

Mathematical and Numerical Methods for Non-linear Beam Dynamics

W. Herr

CERN, Geneva, Switzerland

Abstract

Non-linear effects in accelerator physics are important for both successful operation of accelerators and during the design stage. Since both of these aspects are closely related, they will be treated together in this overview. Some of the most important aspects are well described by methods established in other areas of physics and mathematics. The treatment will be focused on the problems in accelerators used for particle physics experiments. Although the main emphasis will be on accelerator physics issues, some of the aspects of more general interest will be discussed. In particular, we demonstrate that in recent years a framework has been built to handle the complex problems in a consistent form, technically superior and conceptually simpler than the traditional techniques. The need to understand the stability of particle beams has substantially contributed to the development of new techniques and is an important source of examples which can be verified experimentally. Unfortunately, the documentation of these developments is often poor or even unpublished, in many cases only available as lectures or conference proceedings.

1 Introduction and motivation

1.1 Single-particle dynamics

The concepts developed here are used to describe single-particle transverse dynamics in rings, i.e. circular accelerators or storage rings. In the case of linear betatron motion the theory is rather complete and the standard treatment [1] suffices to describe the dynamics. In parallel with this theory the well-known concepts such as closed orbit and Twiss parameters are introduced and emerge automatically from the Courant–Snyder formalism [1]. The formalism and applications are found in many textbooks (e.g. [2,3]).

In many new accelerators or storage rings (e.g. the Large Hadron Collider (LHC)) the description of the machine with a linear formalism becomes insufficient and the linear theory must be extended to treat non-linear effects. The stability and confinement of the particles is not given a priori and should rather emerge from the analysis. Non-linear effects are a main source of performance limitations in such machines. Non-linear dynamics has to be used in many applications and the basics are described in the literature [4,5]. A reliable treatment is required and the progress in recent years allows us to evaluate these effects. A very useful overview and details can be found in [6–8].

In this article, we restrict the treatment to the beam dynamics in rings, since properties such as stability of particles restrict the form of the tools to describe them. However, many of the concepts (e.g. Lie transformations) can be very beneficial for the analysis of single-pass systems.

2 New concepts

The key to the more modern approach shown in this article is to avoid the prejudices about the stability and other properties of the ring. Instead, we must describe the machine in terms of the objects it consists of with all their properties, including the non-linear elements. The analysis will reveal the properties of the particles such as, e.g., stability. In the simplest case, the ring is made of individual machine elements such as magnets which have an existence on their own, i.e. the interaction of a particle with a given element is independent of the motion in the rest of the machine. To successfully study single-particle dynamics, we must be able to describe the action of the machine element on the particle as well as the effect of the machine element in the presence of all other magnets in the ring [9].

2.1 Map-based techniques

In the standard approach to single-particle dynamics in rings, the equations of motion are introduced together with an ansatz to solve these equations. In the case of linear motion, this ansatz is known as the Courant–Snyder treatment. However, this treatment must assume that the motion of a particle in the ring is stable and confined. For a non-linear system this is a priori not known and the attempt to find a complete description of the particle motion must fail.

In a more modern approach, we do not attempt to solve such an overall equation but rather consider the fundamental objects of an accelerator, i.e. the machine elements themselves. These elements, e.g. magnets or other beam elements, are the basic building blocks of the machine. All elements have a well-defined action on a particle which can be described independently of other elements or concepts such as closed orbit or β -functions. Mathematically, they provide a ‘map’ from one face of a building block to the other, i.e. a description of how the particles move inside and between elements. In this context, a map can be anything from linear matrices to high-order integration routines.

The collection of all machine elements makes up the ring and it is the combination of the associated maps which is necessary for the description and analysis of the physical phenomena in the accelerator ring.

The most interesting map is the one which describes the motion once around the machine, the so-called one-turn map (OTM). It contains all necessary information on stability and existence of closed orbit and optical parameters. The reader is assumed to be familiar with this concept in the case of linear beam dynamics [10], where all maps are matrices and the Courant–Snyder analysis of the corresponding OTM produces the desired information such as, e.g., closed orbit or Twiss parameters.

It should therefore be the goal to generalize this concept to non-linear dynamics. The computation of a reliable OTM and the analysis of its properties will provide all relevant information.

2.1.1 Basic concept

Given that the non-linear maps can be rather complex objects, the analysis of the OTM should be separated from the calculation of the map itself.

Using the coordinate vector $\vec{z} = (x, x' = \partial x / \partial s, y, y' = \partial y / \partial s)$, the map \mathcal{M} transforms the coordinates $\vec{z}_1(s_1)$ through a magnet M at position s_1 to new coordinates $\vec{z}_2(s_2)$ at position s_2 :

$$\vec{z}_2(s_2) = \begin{pmatrix} x \\ x' \\ y \\ y' \end{pmatrix}_{s_2} = \mathcal{M} \circ \begin{pmatrix} x \\ x' \\ y \\ y' \end{pmatrix}_{s_1} = \mathcal{M} \circ \vec{z}_1(s_1). \quad (1)$$

The subsequent analysis of the map should reveal the quantities of interest.

3 Linear maps

First I shall introduce the simplest form of maps, i.e. linear maps, and generalize afterwards.

3.1 Examples for linear maps

As a first example, I consider a drift space of length L .

The formal description using the map can be written

$$\begin{pmatrix} x \\ x' \end{pmatrix}_{s_2} = \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix} \circ \begin{pmatrix} x \\ x' \end{pmatrix}_{s_1}. \quad (2)$$

As a second example, I show the map (matrix) for a focusing quadrupole of length L and strength k :

$$\begin{pmatrix} x \\ x' \end{pmatrix}_{s_2} = \begin{pmatrix} \cos L \cdot \sqrt{k} & \frac{1}{\sqrt{k}} \cdot \sin L \cdot \sqrt{k} \\ -\sqrt{k} \cdot \sin L \cdot \sqrt{k} & \cos L \cdot \sqrt{k} \end{pmatrix} \circ \begin{pmatrix} x \\ x' \end{pmatrix}_{s_1}. \quad (3)$$

The map for a quadrupole with short length L (i.e. $1 \gg L \cdot k$) can be approximated by

$$\begin{pmatrix} x \\ x' \end{pmatrix}_{s_2} = \begin{pmatrix} 1 & 0 \\ -k \cdot L (= \frac{1}{f}) & 1 \end{pmatrix} \circ \begin{pmatrix} x \\ x' \end{pmatrix}_{s_1}. \quad (4)$$

These are \mathcal{M} aps, and describe the movement within an element.

3.2 OTMs

Starting from a position s_0 and applying all maps (for N elements) in sequence around a ring with circumference C , we get the OTM for the position s_0 (for one dimension only):

$$\begin{pmatrix} x \\ x' \end{pmatrix}_{s_0+C} = \mathcal{M}_1 \circ \mathcal{M}_2 \circ \dots \circ \mathcal{M}_N \circ \begin{pmatrix} x \\ x' \end{pmatrix}_{s_0} \quad (5)$$

$$\Rightarrow \begin{pmatrix} x \\ x' \end{pmatrix}_{s_0+C} = \mathcal{M}_{\text{ring}}(s_0) \circ \begin{pmatrix} x \\ x' \end{pmatrix}_{s_0}. \quad (6)$$

After this concatenation, we have obtained a OTM for the whole ring (6). In the simplest case all maps are matrices and the concatenation is just the multiplication of linear matrices. Having obtained this OTM, its analysis will give all relevant information; there is no need for any assumptions. To solve Hill's equation using the standard ansatz [10], we need several assumptions such as existence of closed orbit and confinement or periodicity. Having derived the map, we have to extract the information needed. To introduce this topic, we first treat a linear OTM, i.e. a matrix.

3.3 Linear normal forms

The key to the analysis is that maps (here matrices) can be transformed into (Jordan) normal forms. The original maps and the normal forms are equivalent, but the normal form and the required transformation automatically provide the data we need, such as stability, closed orbit, optical parameters (Q , Q' , Twiss function, etc.), invariants of the motion, resonance analysis, and so on.

3.3.1 Linear normal-form transformation

The basic idea is to make a transformation to get a simpler form for the map.

Assuming that the map \mathcal{M}_{12} propagates the variables from location 1 to location 2, we try to find transformations $\mathcal{A}_1, \mathcal{A}_2$ such that

$$\mathcal{A}_1 \mathcal{M}_{12} \mathcal{A}_2^{-1} = \mathcal{R}_{12}. \quad (7)$$

The map \mathcal{R}_{12} is a 'Jordan normal form' (or at least a very simplified form of the map). For example, \mathcal{R}_{12} becomes a pure rotation, which is a simple and typical normal form. The map \mathcal{R}_{12} describes the same dynamics as \mathcal{M}_{12} , but all coordinates are transformed by \mathcal{A}_1 and \mathcal{A}_2 . The transformations $\mathcal{A}_1, \mathcal{A}_2$ 'analyse' the motion.

We assume that the OTM (here a matrix) $\mathcal{M}(s)$ at the position s is (see e.g. the article on transverse dynamics [10])

$$\mathcal{M}(s) = \begin{pmatrix} \cos(\Delta\mu) + \alpha(s) \sin(\Delta\mu) & \beta(s) \sin(\Delta\mu) \\ -\gamma(s) \sin(\Delta\mu) & \cos(\Delta\mu) - \alpha(s) \sin(\Delta\mu) \end{pmatrix}. \quad (8)$$

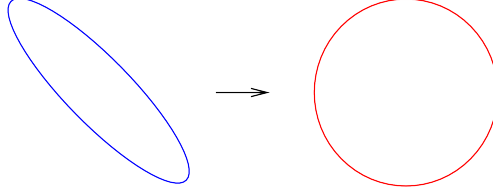


Fig. 1: Normal-form transform for a linear OTM (schematic)

This matrix describes the motion on a phase-space ellipse. We can rewrite \mathcal{M} such that one part \mathcal{R} becomes a pure rotation (a circle), i.e.

$$\mathcal{A}\mathcal{R}\mathcal{A}^{-1} = \mathcal{M}. \quad (9)$$

Remembering lectures on linear algebra, this transformation is done by the evaluation of the eigenvectors and eigenvalues.

Starting from

$$M = \mathcal{A} \circ \mathcal{R}(\Delta\mu) \circ \mathcal{A}^{-1} \quad \text{or} \quad \mathcal{R}(\Delta\mu) = \mathcal{A}^{-1} \circ M \circ \mathcal{A}, \quad (10)$$

we get for the transformation (\mathcal{A}) and the simple form (\mathcal{R})

$$\mathcal{A} = \begin{pmatrix} \sqrt{\beta(s)} & 0 \\ -\frac{\alpha}{\sqrt{\beta}} & \frac{1}{\sqrt{\beta(s)}} \end{pmatrix} \quad \text{and} \quad \mathcal{R} = \begin{pmatrix} \cos(\Delta\mu) & \sin(\Delta\mu) \\ -\sin(\Delta\mu) & \cos(\Delta\mu) \end{pmatrix}. \quad (11)$$

This is just the Courant–Snyder transformation to get β , α , etc. and $\Delta\mu$ is the tune. That is, the Courant–Snyder analysis is just a normal-form transform of the linear one-turn matrix. This transformation works in more than one dimension using the same formalism. The phase-space ellipse is transformed into a circle, which is the representation of the normal form, i.e. the rotation shown in Fig. 1.

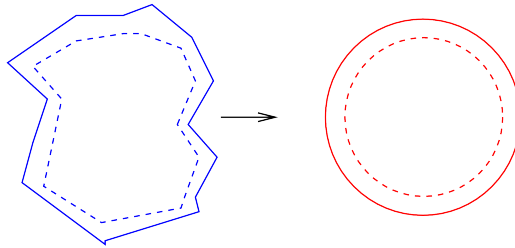


Fig. 2: Normal-form transform for a non-linear OTM (schematic)

Compared with the Courant–Snyder analysis, we have the following properties of this method.

- i) The rotation angle (tune for the OTM) $\Delta\mu$ is contained in the *normalized map*, and we have stability for real values of this phase advance $\Delta\mu$.
- ii) Optical functions (β , α , etc.) are in the *normalizing map*.
- iii) No need to make any assumptions, ansatz or approximation.

In the case of a more complicated OTM (e.g. Fig. 2), the transformation may be more complicated and we have to expect that the rotation angle depends on the amplitude. This is the case for non-linear OTMs, which will be treated in a later section. It has already been mentioned here to introduce the normal-form transformation as a general concept.

3.3.2 Normalized variables

Please note that

$$\begin{pmatrix} x_n \\ x'_n \end{pmatrix} = \mathcal{A}^{-1} \circ \begin{pmatrix} x \\ x' \end{pmatrix} \quad (12)$$

in this simple case is just a variable transformation to new, normalized variables.

3.3.3 Action-angle variables

It is common practice to define action-angle variables. Once the particles ‘travel’ on a circle, we can introduce new variables for simplification and easier interpretation of the results.

- i) Radius (say: $\sqrt{2J}$, with $J = \frac{x_n^2 + x_n'^2}{2}$) is constant (invariant of motion): **action** J .
- ii) Phase advances by constant amount: **angle** Ψ .

The invariant of motion J should be the same at all locations in the ring and is separate from the angle ψ . For the propagation of coordinates around the ring, both are relevant. This separation of the invariant from the angle again becomes important in a later section.

3.3.4 Coupling

The formalism can easily be used to study the case of coupling between the two transverse planes, with the map (in two dimensions)

$$T = \begin{pmatrix} M & n \\ m & N \end{pmatrix}, \quad (13)$$

where M, m, N, n are 2×2 matrices. In the case of coupling, $m \neq 0, n \neq 0$.

We can try to rewrite this using the normal-form algorithms, as before:

$$T = \begin{pmatrix} M & n \\ m & N \end{pmatrix} = V R V^{-1}, \quad (14)$$

with

$$R = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} \gamma I & C \\ -C^t & \gamma I \end{pmatrix}. \quad (15)$$

What have we obtained? The matrix R is our simple rotation in two planes and we have in addition:

- i) A and B are the one-turn matrices for the ‘normal modes’;
- ii) the matrix V transforms from the coordinates (x, x', y, y') into the ‘normal-mode’ coordinates (w, w', v, v') via the expression $(x, x', y, y') = V(w, w', v, v')$;
- iii) the matrix C contains the ‘coupling coefficients’.

3.3.5 Multiple turns

Normal forms are extremely useful when a map should be applied k times (e.g. k turns): the brute-force method would be to multiply the OTM k times, which can be very elaborate and expensive in computing time. Instead, we can first make a transformation into a normal form with a rotation matrix R , which contains the phase advance per turn μ . For k turns, we have to simply replace the phase advance μ by the phase advance for k turns, which is of course $k \cdot \mu$ (16).

$$M^k(x, x') = A R^k A^{-1}(x, x') = A R^k(X, X'). \quad (16)$$

Transforming the rotation back into the usual form returns the k -turn map.

4 Non-linear maps

In this section, I should like to introduce *non-linear maps*. I shall start with a linear map and then generalize. The effect of a (short) quadrupole depends linearly on amplitude as (rewritten from the matrix form)

$$\vec{z}(s_2) = \begin{pmatrix} x \\ x' \\ y \\ y' \end{pmatrix}_{s_2} = \begin{pmatrix} x \\ x' \\ y \\ y' \end{pmatrix}_{s_1} + \begin{pmatrix} 0 \\ k_1 \cdot x_{s_1} \\ 0 \\ k_1 \cdot y_{s_1} \end{pmatrix}. \quad (17)$$

Or, in the already known convention using the matrix \mathbf{M} , $\vec{z}(s_2) = \mathbf{M} \cdot \vec{z}(s_1)$.

4.1 Symplecticity

A key concept for the application of maps is a requirement called ‘symplecticity’. In a ring, not all possible maps are allowed. This requires for a matrix \mathcal{M} that it fulfils the condition

$$\mathcal{M}^T \cdot S \cdot \mathcal{M} = S \quad (18)$$

with

$$S = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}. \quad (19)$$

The physical meaning of this requirement is that \mathcal{M} is area preserving and

$$\lim_{n \rightarrow \infty} \mathcal{M}^n = \text{finite} \implies \det \mathcal{M} = 1. \quad (20)$$

It should be noted that $\det \mathcal{M} = 1$ is required, but not sufficient, for \mathcal{M} to be symplectic.

4.2 Taylor maps

The effect of a (thin) sextupole with strength k_2 is in the form of (17)

$$\vec{z}(s_2) = \begin{pmatrix} x \\ x' \\ y \\ y' \end{pmatrix}_{s_2} = \begin{pmatrix} x \\ x' \\ y \\ y' \end{pmatrix}_{s_1} + \begin{pmatrix} 0 \\ \frac{1}{2}k_2 \cdot (x_{s_1}^2 - y_{s_1}^2) \\ 0 \\ k_2 \cdot (x_{s_1} \cdot y_{s_1}) \end{pmatrix}. \quad (21)$$

We can write it now formally as

$$\vec{z}(s_2) = \mathcal{M} \circ \vec{z}(s_1). \quad (22)$$

Here \mathcal{M} is not a matrix but a non-linear map. It cannot be written in linear matrix form. We need something like (for the vector element x)

$$\begin{aligned} \vec{z}_x(s_2) = x(s_2) &= R_{11} \cdot x + R_{12} \cdot x' + R_{13} \cdot y + \dots \\ &+ T_{111} \cdot x^2 + T_{112} \cdot x x' T_{122} \cdot x'^2 + \\ &+ T_{113} \cdot x y + T_{114} \cdot x y' + \dots \\ &+ U_{1111} \cdot x^3 + U_{1112} \cdot x^2 x' + \dots \end{aligned} \quad (23)$$

and the equivalent for other variables. Please note that $x(s_2)$ now depends also on the second transverse variable y .

4.2.1 Forms of Taylor maps

We can write (23) for $j = 1, \dots, 4$:

$$z_j(s_2) = \sum_{k=1}^4 R_{jk} z_k(s_1) + \sum_{k=1}^4 \sum_{l=1}^4 T_{jkl} z_k(s_1) z_l(s_1). \quad (24)$$

If this map is truncated after the second order, we call it a second-order map and formally write it as

$$\mathcal{A}_2 = [R, T]. \quad (25)$$

Higher orders can be defined as needed:

$$\mathcal{A}_3 = [R, T, U] \implies [R, T] + \sum_{k=1}^4 \sum_{l=1}^4 \sum_{m=1}^4 U_{jklm} z_k(s_1) z_l(s_1) z_m(s_1), \quad (26)$$

which is truncated after the third order.

As an example, the complete second-order map for a (thick) sextupole with length L and strength K (in two dimensions) is

$$\begin{aligned} x_2 &= x_1 + Lx'_1 - K \left(\frac{L^2}{4}(x_1^2 - y_1^2) + \frac{L^3}{6}(x_1 x'_1 - y_1 y'_1) + \frac{L^4}{24}(x_1'^2 - y_1'^2) \right), \\ x'_2 &= x'_1 - K \left(\frac{L}{2}(x_1^2 - y_1^2) + \frac{L^2}{2}(x_1 x'_1 - y_1 y'_1) + \frac{L^3}{6}(x_1'^2 - y_1'^2) \right), \\ y_2 &= y_1 + Ly'_1 + K \left(\frac{L^2}{4}x_1 y_1 + \frac{L^3}{6}(x_1 y'_1 + y_1 x'_1) + \frac{L^4}{24}(x_1' y_1') \right), \\ y'_2 &= y'_1 + K \left(\frac{L}{2}x_1 y_1 + \frac{L^2}{2}(x_1 y'_1 + y_1 x'_1) + \frac{L^3}{6}(x_1' y_1') \right). \end{aligned} \quad (27)$$

The definition of K is not unique; it can differ by some factor, e.g.

$$\left(\frac{\partial^2 x}{\partial t^2} = k \cdot x^2 \quad \text{versus} \quad \frac{\partial^2 x}{\partial t^2} = \frac{k}{2} \cdot x^2 \right). \quad (28)$$

4.2.2 Symplecticity of Taylor maps

We have argued that a vital requirement for maps is the symplecticity. This must hold also for non-linear maps; however:

- i) truncated Taylor expansions are not matrices;
- ii) it is the associated Jacobian matrix \mathcal{J} which must fulfil the symplecticity condition (see (18) and (19)):

$$\mathcal{J}_{ik} = \frac{\partial z_2^i}{\partial z_1^k} \quad \left(\text{e.g. } \mathcal{J}_{xy} = \frac{\partial z_2^x}{\partial z_1^y} \right), \quad (29)$$

where \mathcal{J} must fulfil $\mathcal{J}^t \cdot S \cdot \mathcal{J} = S$.

In general, $\mathcal{J}_{ik} \neq \text{const} \rightarrow$ for a truncated Taylor map; this can be difficult to fulfil for all z .

As an example, we take the sextupole map (for simplicity in one dimension) [21]

$$\begin{aligned} x_2 &= x_1 + Lx'_1 - K \left(\frac{L^2}{4}x_1^2 + \frac{L^3}{6}x_1 x'_1 + \frac{L^4}{24}x_1'^2 + \mathcal{O}(3) \right), \\ x'_2 &= x'_1 - K \left(\frac{L}{2}x_1^2 + \frac{L^2}{2}x_1 x'_1 + \frac{L^3}{6}x_1'^2 + \mathcal{O}(3) \right). \end{aligned} \quad (30)$$

Then we can compute

$$\mathcal{J}^T \cdot S \cdot \mathcal{J} = \begin{pmatrix} 0 & 1+\Delta S \\ -1-\Delta S & 0 \end{pmatrix} \neq S. \quad (31)$$

We find that \mathcal{J} is non-symplectic with the error

$$\Delta S = \frac{K}{72} L^4 (L^2 x'^2 + 6Lx x' + 6x^2). \quad (32)$$

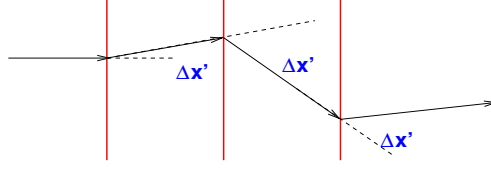


Fig. 3: Movement of a particle through thin elements

5 Thin elements

From Eq. (32), we may get a hint of how to attack this problem. The error ΔS compared to the fully symplectic map depends on the fourth power of the sextupole length L^4 . This can be extended to other elements as well and for short objects the error may be negligible. We can attempt to ‘make’ all elements thin, i.e. short enough to have a negligible non-symplecticity. For this purpose, we define thin elements, i.e. an element with *no* length, because they avoid problems with symplecticity of non-linear maps such as Taylor maps. The error in equation (32) becomes zero.

5.1 Concept of thin elements

Real magnets have a finite length, i.e. they are thick magnets. The field and length are used to compute their effect, i.e. the map; in the simplest case it becomes a matrix (3). In passing, we have already mentioned the equivalent matrix for a quadrupole with zero length (4), since it is often used in analytical calculations. A severe consequence of thick elements is that they are not always linear elements (even not dipoles or quadrupoles). For thick, non-linear magnets a closed solution for maps often does not exist.

The procedure we shall apply to get ‘thin’ magnets is to let the length go to zero, but keep the field integral finite (constant). As consequences, thin dipoles and quadrupoles are linear elements and all thin magnets are much easier to use and automatically symplectic. The movement of a particle through thin elements is shown schematically in Fig. 3. The key is that there is no change of amplitudes x and y at the element; only the momenta x' and y' receive an amplitude-dependent deflection (kick)

$$x' \rightarrow x' + \Delta x' \quad \text{and} \quad y' \rightarrow y' + \Delta y'. \quad (33)$$

Modelling a magnet with a finite length by a thin magnet, we have to pay a price and must ask under which conditions this is a good strategy. We should expect that this approximation is valid when the finite length is small or when the length does not matter. The main questions to answer are as follows.

- i) What happens to the accuracy and the implications for the beam dynamics?
- ii) What is the best implementation of thin magnets to minimize the side effects?

We should study these issues with some simple examples and generalize later.

5.2 Accuracy of thin elements

We start again with the exact matrix for a thick quadrupole with the strength K and the length L .

$$\mathcal{M}_{s \rightarrow s+L} = \begin{pmatrix} \cos L \cdot \sqrt{K} & \frac{1}{\sqrt{K}} \cdot \sin L \cdot \sqrt{K} \\ -\sqrt{K} \cdot \sin L \cdot \sqrt{K} & \cos L \cdot \sqrt{K} \end{pmatrix}. \quad (34)$$

This matrix describes the motion in a thick quadrupole exactly and we can estimate the errors we make by our approximation. Although this demonstration is done for a quadrupole, the final result is valid for all linear and non-linear elements.

From the exact matrix

$$\mathcal{M}_{s \rightarrow s+L} = \begin{pmatrix} \cos L \cdot \sqrt{K} & \frac{1}{\sqrt{K}} \cdot \sin L \cdot \sqrt{K} \\ -\sqrt{K} \cdot \sin L \cdot \sqrt{K} & \cos L \cdot \sqrt{K} \end{pmatrix}, \quad (35)$$

we make a Taylor expansion in ‘small’ length L :

$$L^0 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + L^1 \cdot \begin{pmatrix} 0 & 1 \\ -K & 0 \end{pmatrix} + L^2 \cdot \begin{pmatrix} -\frac{K}{2} & 0 \\ 0 & -\frac{K}{2} \end{pmatrix} + \dots \quad (36)$$

and keep up to the *first-order* term in L :

$$\mathcal{M}_{s \rightarrow s+L} = L^0 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + L^1 \cdot \begin{pmatrix} 0 & 1 \\ -K & 0 \end{pmatrix}, \quad (37)$$

$$\mathcal{M}_{s \rightarrow s+L} = \begin{pmatrix} 1 & L \\ -K \cdot L & 1 \end{pmatrix} + \mathcal{O}(L^2). \quad (38)$$

We make two observations from the approximation (38).

- i) It is precise to first order $\mathcal{O}(L^1)$ and the error is to second order $\mathcal{O}(L^2)$.
- ii) We have $\det \mathcal{M} \neq 1$; therefore, it cannot be symplectic (see 18 and 19).

To ensure the symplecticity, we need to modify the matrix. We can try to add a term in the matrix:

$$\mathcal{M}_{s \rightarrow s+L} = \begin{pmatrix} 1 & L \\ -K \cdot L & 1 \end{pmatrix} + \mathcal{O}(L^2) \quad (39)$$

and we obtain

$$\mathcal{M}_{s \rightarrow s+L} = \begin{pmatrix} 1 & L \\ -K \cdot L & 1 - KL^2 \end{pmatrix}. \quad (40)$$

The matrix (40) is still precise to first order $\mathcal{O}(L^1)$ but ‘symplectified’ by adding the term $-KL^2$. The added term is inaccurate to second order (because of L^2) and therefore the order of accuracy is not changed by this modification. This procedure restores the symplecticity, but does not change the level of accuracy.

For the next step, we keep up to the *second order* in L from the Taylor expansion (36):

$$\mathcal{M}_{s \rightarrow s+L} = \begin{pmatrix} 1 - \frac{1}{2}KL^2 & L \\ -K \cdot L & 1 - \frac{1}{2}KL^2 \end{pmatrix} + \mathcal{O}(L^3). \quad (41)$$

It is precise to second order $\mathcal{O}(L^2)$ and therefore more accurate than (40), but again not symplectic.

We proceed as before and the symplectification now looks like

$$\mathcal{M}_{s \rightarrow s+L} = \begin{pmatrix} 1 - \frac{1}{2}KL^2 & L - \frac{1}{4}KL^3 \\ -K \cdot L & 1 - \frac{1}{2}KL^2 \end{pmatrix} + \mathcal{O}(L^3). \quad (42)$$

The accuracy is maintained to second order $\mathcal{O}(L^2)$ and it is now fully symplectic.

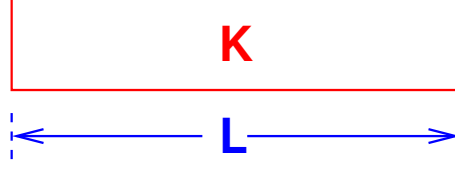


Fig. 4: Schematic picture of a quadrupole with length L and strength K (convention)

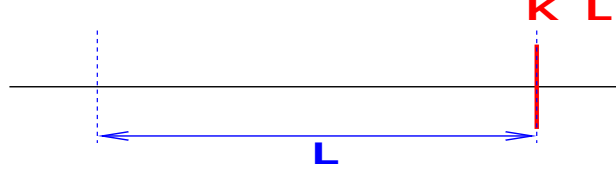


Fig. 5: Schematic picture of a quadrupole with length L and a kick with strength K at the end of the element

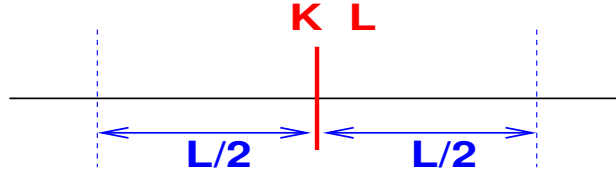


Fig. 6: Schematic picture of a quadrupole with length L and a kick with strength K in the centre of the element

5.2.1 Physical meaning of the symplectification

In principle, one could go on with such a scheme, but what is a good strategy? The main issues for using an implementation are accuracy and speed. First, we have a look at the significance of the symplectification, which looks rather arbitrary. Assume a linear element (quadrupole) of length L and strength K as illustrated in (Fig. 4). A very simple way to apply a thin-lens kick is shown in (Fig. 5). The quadrupole is treated as a drift length of length L followed by a thin-length kick with the strength $K \cdot L$.

We can compute the full matrix of the drift and the kick and obtain

$$\begin{pmatrix} 1 & 0 \\ -K \cdot L & 1 \end{pmatrix} \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & L \\ -K \cdot L & 1 - KL^2 \end{pmatrix}. \quad (43)$$

We find that this resembles the ‘symplectification’ we applied to the truncated map of order $\mathcal{O}(L^1)$. Because of the symmetry, we would get the same result when we apply the kick before the drift.

Another option is to apply the thin-lens kick in the centre of the element, as shown in Fig. 6. This time we apply the kick in the centre of the element, preceded and followed by drift spaces of the length $L/2$.

The multiplication of the three matrices leads us to

$$\mathcal{M}_{s \rightarrow s+L} = \begin{pmatrix} 1 - \frac{1}{2}KL^2 & L - \frac{1}{4}KL^3 \\ -K \cdot L & 1 - \frac{1}{2}KL^2 \end{pmatrix} + \mathcal{O}(L^3). \quad (44)$$

This is equivalent to the truncated and symplectified map of order $\mathcal{O}(L^2)$.

We can summarize the two options as follows.

- i) One kick at the end or entry: error (inaccuracy) is of first order $\mathcal{O}(L^1)$.
- ii) One kick in the centre: error (inaccuracy) is of second order $\mathcal{O}(L^2)$.

We find that it is very relevant how to apply thin lenses and the aim should be to be precise and fast (and simple to implement, for example in a computer program).

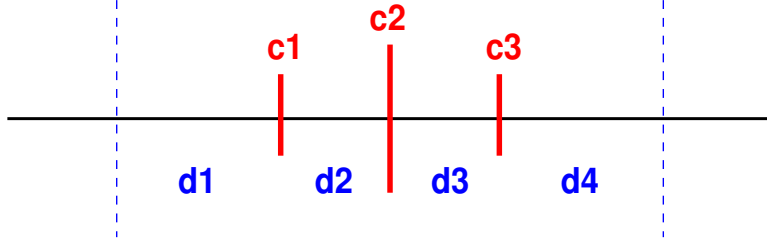


Fig. 7: Symplectic integrator with three thin lenses (schematic)

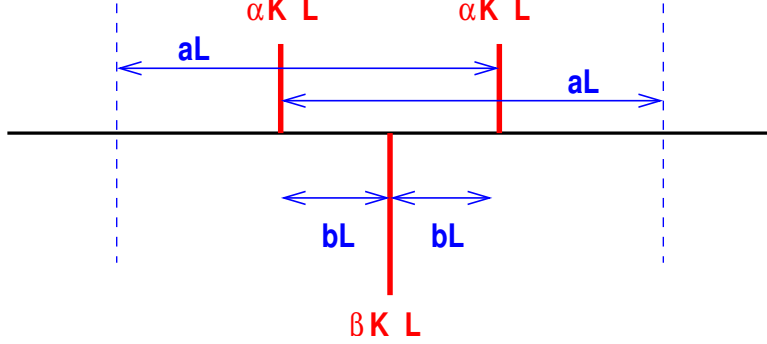


Fig. 8: Symplectic integrator with three thin lenses (schematic)

5.3 Symplectic integration

We introduce now the concept of symplectic integration and start with simple examples and generalize the concept afterwards. Different integration methods exist which yield symplectic maps and we shall discuss some of them which are relevant for our discussion of applications to accelerators.

5.3.1 Can we do better than in the previous examples?

We can try a model with three kicks [18], as shown in Fig. 7. We assume that the kicks ($c1$, $c2$, $c3$) and drift spaces ($d1$, $d2$, $d3$, $d4$) are variable. We can look for an optimization of these variables to get the best accuracy (i.e. smallest deviation from the exact solution). Since we have only thin lenses, the symplecticity is automatically ensured.

The results of the optimization of the strengths and the drift spaces between thin kicks are shown in Fig. 8. The optimization of the free parameters gives the values (see Fig. 8): $a \approx 0.6756$, $b \approx -0.1756$, $\alpha \approx 1.3512$, $\beta \approx -1.7024$.

It can easily be shown [18] that using this technique with three kicks, we obtain a $\mathcal{O}(4)$ integrator. We should expect that more kicks improve the accuracy and as an example a $\mathcal{O}(6)$ integrator would require nine kicks. This will be shown next.

5.3.2 Yoshida formalism

What we have done is a symplectic integration [18, 19]. From a lower order integration scheme (one kick), we can construct a higher order scheme. Formally (for the formulation of $S_k(t)$ see later), we can write: from a second-order scheme (one kick) $S_2(t)$ we construct a fourth-order scheme (three kicks = 3×1 kick) like

$$S_4(t) = S_2(x_1 t) \circ S_2(x_0 t) \circ S_2(x_1 t) \quad (45)$$

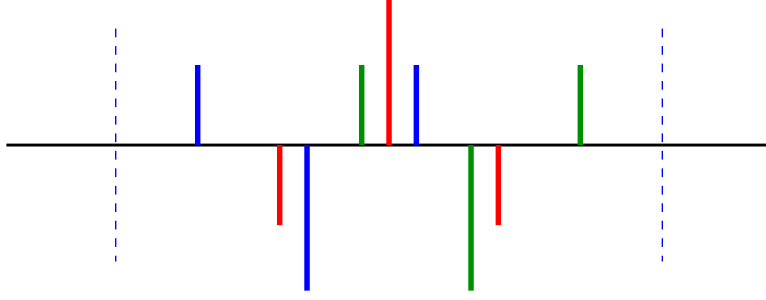


Fig. 9: Symplectic integrator with nine thin lenses (schematic)

with the coefficients

$$x_0 = \frac{-2^{1/3}}{2 - 2^{1/3}} \approx -1.7024, \quad x_1 = \frac{1}{2 - 2^{1/3}} \approx 1.3512. \quad (46)$$

Equation (45) should be understood symbolically in that we construct a fourth-order integrator from three second-order integrators. As a further example from a fourth order to a sixth order, we have to use

$$S_6(t) = S_4(x_1 t) \circ S_4(x_0 t) \circ S_4(x_1 t). \quad (47)$$

The interpretation of this construct is that we get three times fourth order with three kicks each; we have the nine-kick, sixth-order integrator mentioned earlier. This is shown schematically in Fig. 9. The three different colours correspond to the three fourth-order integrators, put together into a sixth-order integrator. The sixth-order integrator requires nine kicks and we have three interleaved fourth-order integrators. Going to higher order such as eighth order, we need three interleaved sixth-order integrators, altogether 27 thin lenses. The key is that this can be used as an iterative scheme to go to very high orders and can easily be implemented on a computer. More importantly, the coefficients in (46) are the same in this iterative procedure.

As a general formula, we can express this procedure as

$$S_{k+2}(t) = S_k(x_1 t) \circ S_k(x_0 t) \circ S_k(x_1 t). \quad (48)$$

In practice, this means that each kick in the k th-order integrator is replaced by a k th-order integrator scaled accordingly to obtain a scheme of $(k + 2)$ th order.

We have used a linear map (quadrupole) to demonstrate the integration, but it can be applied to other maps (solenoids, higher-order, non-linear maps). It was shown [18, 19] that this is indeed possible and, most importantly, we get the same integrators. The proof and the systematic extension can be done in the form of Lie operators. It can be proven that this scheme gives the theoretically best possible accuracy for a given number of thin elements.

5.3.3 Integration of non-linear elements

We demonstrate now this formalism for non-linear elements. Let us assume a general case:

$$x'' = f(x). \quad (49)$$

The disadvantage of a kick (49) is that usually a closed solution through the element does not exist and an integration is necessary.

The advantage of non-linear elements is that they are usually thin (at least thinner than dipoles or quadrupoles). As examples, we give the lengths of typical LHC magnets.

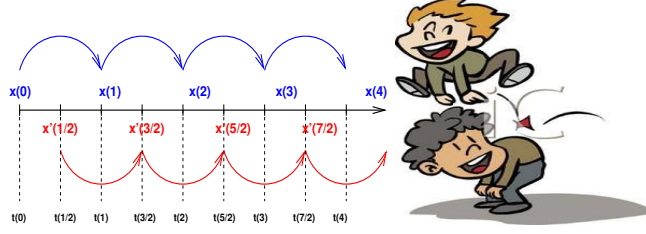


Fig. 10: Schematic representation of a ‘leap-frog’ integration scheme

- i) Dipoles: ≈ 14.3 m.
- ii) Quadrupoles: $\approx 2\text{--}5$ m.
- iii) Sextupoles, octupoles: ≈ 0.30 m.

We can hope to obtain a good approximation for elements with small lengths. To get an estimate, we try our simplest $\mathcal{O}(2)$ thin-lens approximation first. Therefore, we represent the non-linear element as a thin-lens kick (49) with drift spaces of $L/2$ before and after the kick.

5.3.4 Accuracy of thin lenses: our $\mathcal{O}(2)$ model

Assuming the general case for a non-linear kick $\Delta x'$,

$$x'' = f(x) \quad (= \Delta x'), \quad (50)$$

we follow the three steps:

$$\text{Step 1.} \quad \begin{pmatrix} x \\ x' \end{pmatrix}_{s_1+L/2} = \begin{pmatrix} 1 & L/2 \\ 0 & 1 \end{pmatrix} \circ \begin{pmatrix} x \\ x' \end{pmatrix}_{s_1}; \quad (51)$$

$$\text{Step 2.} \quad \begin{pmatrix} x \\ x' \end{pmatrix}_{s_1+L/2} = \begin{pmatrix} x \\ x' + \Delta x' \end{pmatrix}_{s_1+L/2}; \quad (52)$$

$$\text{Step 3.} \quad \begin{pmatrix} x \\ x' \end{pmatrix}_{s_1+L} = \begin{pmatrix} 1 & L/2 \\ 0 & 1 \end{pmatrix} \circ \begin{pmatrix} x \\ x' \end{pmatrix}_{s_1+L/2}. \quad (53)$$

Using this thin-lens approximation ($\mathcal{O}(2)$) gives us for the map

$$\begin{aligned} x'(L) &\approx x'_0 + Lf(x_0 + \frac{L}{2}x'_0), \\ x(L) &\approx x_0 + \frac{L}{2}(x'_0 + x'(L)). \end{aligned} \quad (54)$$

Here we write x_0 and x'_0 for the coordinates and angles at the entry of the element. The result (54) corresponds to the well-known ‘leap-frog’ algorithm/integration, schematically shown in Fig. 10. It is symplectic and time reversible.

For any $x'' = f(x, x', t)$, we can solve it using

$$\begin{aligned} x'_{i+3/2} &\approx x'_{i+1/2} + f(x_{i+1})\Delta t, \\ x_{i+1} &\approx x_i + x'_{i+1/2}\Delta t, \end{aligned} \quad (55)$$

which is what is known as a ‘leap-frog’ integration.

5.3.5 Accuracy of the ‘leap-frog’ algorithm/integration

The (exact) Taylor expansion gives

$$x(L) = x_0 + x'_0 L + \frac{1}{2} f(x_0) L^2 + \frac{1}{6} x'_0 f'(x_0) L^3 + \dots; \quad (56)$$

the (approximate) ‘leap-frog’ algorithm (55) gives

$$x(L) = x_0 + x'_0 L + \frac{1}{2} f(x_0) L^2 + \frac{1}{4} x'_0 f'(x_0) L^3 + \dots. \quad (57)$$

The two methods differ from L^3 onwards and we have an integrator of order $\mathcal{O}(L^3)$. This had to be expected since we know from previous considerations that this type of splitting gives such an accuracy. For small L , this method is acceptable and symplectic.

As an application, we assume a (one-dimensional) sextupole with (definition of k not unique)

$$x'' = k \cdot x^2 = f(x); \quad (58)$$

using the thin-lens approximation using (54) gives

$$\begin{aligned} x(L) &= x_0 + x'_0 L + \frac{1}{2} k x_0^2 L^2 + \frac{1}{2} k x_0 x'_0 L^3 + \frac{1}{8} k x_0'^2 L^4, \\ x'(L) &= x'_0 + k x_0^2 L + k x_0 x'_0 L^2 + \frac{1}{4} k x_0'^2 L^3. \end{aligned} \quad (59)$$

This is the map for a thick sextupole of length L in thin-lens approximation, accurate to $\mathcal{O}(L^2)$.

6 Hamiltonian formalism

In this section, we introduce the Hamiltonian formalism for the classical motion of a particle [20]. We describe the motion of a system by a function L :

$$L(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t), \quad (60)$$

where (q_1, \dots, q_n) are generalized coordinates and $(\dot{q}_1, \dots, \dot{q}_n)$ are generalized velocities. For the introduction of the concepts, we use q for the coordinate, since it is commonly used in the literature [20]. The function L defines the Lagrange function and the integral

$$I = \int L(q_i, \dot{q}_i, t) dt \quad (61)$$

defines the action. Without proof or derivation, we quote [20]

$$L = T - V, \quad (62)$$

where T is the kinetic energy and V the potential energy.

The Hamiltonian principle states that a system moves such that the action I becomes an extremum when different paths are possible. This is schematically shown in Fig. 11. The action becomes a minimum for the red path. Such concepts are known from other applications such as optics or quantum mechanics using Feynman path integrals.

Without proof, the action

$$I = \int_1^2 L(q_i, \dot{q}_i, t) dt = \text{extremum} \quad (63)$$

becomes an extremum (minimum) when the Lagrange function fulfils the relation [20]

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0, \quad (64)$$

which is usually called the Euler–Lagrange equation.

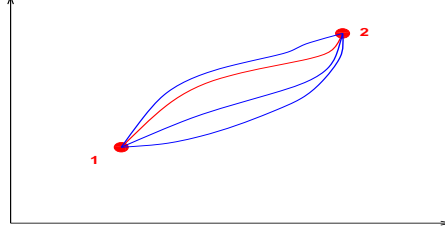


Fig. 11: Schematic of the Hamiltonian principle. Different paths and path with minimum action

6.1 From the Lagrangian to the Hamiltonian

The Lagrangian $L(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n, t)$ in generalized coordinates and velocities provides (n) second-order differential equations. We can try to get first-order differential equations (which are always solvable) and generalized momenta instead of velocities. This leads us to the Hamiltonian function and the Hamiltonian formalism. The main differences are as follows.

- i) Lagrangian:
 - a) n second-order equations;
 - b) n -dimensional coordinate space.
- ii) Hamiltonian:
 - a) $2n$ first-order equations;
 - b) $2n$ -dimensional phase space.

Corresponding (so-called conjugate) momenta p_j are

$$p_j = \frac{\partial L}{\partial \dot{q}_j}. \quad (65)$$

Once we know what the canonical momenta p_i are, the Hamiltonian is a transformation of the Lagrangian:

$$H(q_j, p_j, t) = \sum_i \dot{q}_i p_i - L(q_j, \dot{q}_j, t). \quad (66)$$

Without proof or derivation, we quote [20]

$$H = T + V, \quad (67)$$

where T is the kinetic energy and V the potential energy.

From the Hamiltonian principle, we obtain two first-order equations of motion:

$$\frac{\partial H}{\partial q_j} = -\dot{p}_j = -\frac{dp_j}{dt}, \quad \frac{\partial H}{\partial p_j} = \dot{q}_j = \frac{dq_j}{dt}. \quad (68)$$

The key difference to the Lagrangian is that the canonical momenta can be considered as independent variables (contrary to velocities of the coordinates) and are on equal footing with the coordinates. They form a separate set of variables and their time development is treated separately (68).

6.2 Hamiltonian of electromagnetic elements in accelerators

For the Hamiltonian of a (relativistic) particle in an electromagnetic field, we have (using from now on the more familiar variable x as the coordinate)

$$\mathcal{H}(\vec{x}, \vec{p}, t) = c\sqrt{(\vec{p} - e\vec{A}(\vec{x}, t))^2 + m_0^2 c^2} + e\Phi(\vec{x}, t), \quad (69)$$

where $\vec{A}(\vec{x}, t)$, $\Phi(\vec{x}, t)$ are the vector and scalar potentials. Using canonical variables and the design path length s as independent variable (bending in x -plane),

$$\mathcal{H} = -\left(1 + \frac{x}{\rho}\right) \cdot \sqrt{(1 + \delta)^2 - p_x^2 - p_y^2} + \frac{x}{\rho} + \frac{x^2}{2\rho^2} - \frac{A_s(x, y)}{B_0\rho}, \quad (70)$$

where $\delta = (p - p_0)/p$ is the relative momentum deviation and $A_s(x, y)$ the longitudinal component of the vector potential.

In general, the magnetic fields can be described with the multipole expansion

$$B_y + iB_x = \sum_{n=1} (b_n + ia_n)(x + iy)^{n-1} \quad (71)$$

and, since $\vec{B} = \nabla \times \vec{A}$ for the vector potential \vec{A} ,

$$A_s(x, y) = \sum_{n=1} \frac{1}{n} [(b_n + ia_n)(x + iy)^n]. \quad (72)$$

Here $n = 1$ refers to dipolar fields (not always the case and different conventions exist). For a large machine ($x \ll \rho$), we expand the root and sort the variables to get

$$\mathcal{H} = \overbrace{\frac{p_x^2 + p_y^2}{2(1 + \delta)}}^{\text{kinematic}} - \underbrace{\frac{x\delta}{\rho}}_{\text{dispersive}} + \underbrace{\frac{x^2}{2\rho^2}}_{\text{focusing}} + \overbrace{\frac{k_1}{2}(x^2 - y^2)}^{\text{quadrupole}} + \overbrace{\frac{k_2}{6}(x^3 - 3xy^2)}^{\text{sextupole}} \quad (73)$$

$$\left(\text{using MAD convention: } k_n = \frac{1}{B\rho} \frac{\partial^n B_y}{\partial x^n} \right). \quad (74)$$

We can see the contribution to the different multipole fields. The Hamiltonian describes *exactly* the motion of a particle *through a magnet*. This is our basis to extend the linear to a non-linear formalism.

But how do we apply it to an accelerator? A standard method used in the past was to derive a grandiose Hamiltonian for the whole ring. In that case one views the machine elements (magnets) as a localized fluctuation in the variable s , requiring a large set of Fourier modes. In this traditional treatment one tries to find a global description of the entire ring to derive the aspects of beam dynamics of interest. This goal is usually doomed to fail for complex systems, in particular with localized strong and non-linear objects such as beam-beam interactions in a collider. It is the humble opinion of the author that such an approach does not reach beyond (doubtful) pedagogical purposes.

Instead, I shall treat individual magnets as objects having an existence on their own. The magnets are treated as *local* objects where we can compute the map for the magnet. In a global view it is not possible to have an idea of the effect of a single magnet (e.g. a quadrupole) unless the Twiss parameters at the locations of the magnets are known. This requires a priori knowledge of the OTM.

This ‘local’ approach is clearly influenced by the computer simulation of accelerators where conceptually all elements are treated as separate objects, which are described (and used) independently of

‘global’ concepts such as closed orbits, beta functions etc. With this treatment the properties of the ring such as stability and optical functions will arise naturally, as demonstrated in the simple introduction to linear normal forms. The description of machine elements can be easily achieved by defining the maps through their individual Hamiltonians. This allows naturally the treatment of non-linear elements and in connection to simulations it translates itself into the ‘object-oriented’ approach in computer science. Concepts built on this approach are described in detail in [6].

In the next step we proceed to present the necessary tools.

6.3 Poisson brackets

We introduce a Poisson bracket for a differential operator [20]:

$$[f, g] = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial x_i} \right). \quad (75)$$

Here the variables x_i, p_i are canonical variables; f and g are any functions of x_i and p_i .

We can now write (using the Hamiltonian H for the function g in (75))

$$f(x_i, p_i) = x_i \Rightarrow [x_i, H] = \frac{\partial H}{\partial p_i} = \frac{dx_i}{dt}, \quad (76)$$

$$f(x_i, p_i) = p_i \Rightarrow [p_i, H] = -\frac{\partial H}{\partial x_i} = \frac{dp_i}{dt}. \quad (77)$$

Therefore, Poisson brackets encode Hamilton’s equations [20]. In general, the Poisson bracket $[f, H]$ describes the time evolution of the system (i.e. the function f). It is a special case of

$$\frac{df}{dt} = [f, H] + \frac{\partial f}{\partial t}. \quad (78)$$

If H does not explicitly depend on time, the condition

$$[f, H] = 0 \quad (79)$$

implies that f is an invariant of the motion and therefore the Poisson brackets determine invariants of the motion.

6.4 Lie operators

6.4.1 Definition

We can define

$$: f : g = [f, g], \quad (80)$$

where $: f :$ is an operator acting on the function g :

$$: f := [f,]. \quad (81)$$

The operator $: f :$ is called a Lie operator.

We could say colloquially that Lie operators are Poisson brackets in waiting. A Lie operator acts on functions $g(x, p)$, and we have special cases where the function g itself is a canonical variable:

$$g(x, p) = x \rightarrow : f : x, \quad (82)$$

$$g(x, p) = p \rightarrow : f : p. \quad (83)$$

6.4.2 Useful formulas for calculations

With x the coordinate and p the momentum, we can write useful and frequently used expressions for special cases of f . From (75) and (81), it follows that

$$:x:=\frac{\partial}{\partial p}, \quad :p:=-\frac{\partial}{\partial x}, \quad (84)$$

$$:x:^2=\frac{\partial^2}{\partial p^2}, \quad :p:^2=\frac{\partial^2}{\partial x^2}, \quad (85)$$

$$:x^2:=2x\frac{\partial}{\partial p}, \quad :p^2:=-2p\frac{\partial}{\partial x}, \quad (86)$$

$$:xp:=p\frac{\partial}{\partial p}-x\frac{\partial}{\partial x}, \quad :x::p:=:p::x:=-\frac{\partial^2}{\partial x\partial p}. \quad (87)$$

Applying those to the coordinate x and the momentum p , we find for later use

$$:p^2:x=\frac{\partial p^2}{\partial x}\frac{\partial x}{\partial p}-\frac{\partial p^2}{\partial p}\frac{\partial x}{\partial x}=-2p, \quad (88)$$

$$:p^2:p=\frac{\partial p^2}{\partial x}\frac{\partial p}{\partial p}-\frac{\partial p^2}{\partial p}\frac{\partial p}{\partial x}=0, \quad (89)$$

$$(:p^2:)^2x=:p^2: (:p^2:x)=:p^2: (-2p)=0, \quad (90)$$

$$(:p^2:)^2p=:p^2: (:p^2:p)=:p^2: (0)=0. \quad (91)$$

6.5 Lie transformations

We can define powers of Lie operators as

$$(:f:)^2g=:f: (:f:g)=[f,[f,g]], \text{ etc.} \quad (92)$$

and in particular extend this to a series of the form

$$e^{:f:}=\sum_{i=0}^{\infty}\frac{1}{i!}(:f:)^i, \quad (93)$$

$$e^{:f:}=1+:f: +\frac{1}{2!}(:f:)^2+\frac{1}{3!}(:f:)^3+\dots. \quad (94)$$

The operator $e^{:f:}$ is called a Lie transformation and f the generator of the transformation. The generator f can be any function of x and p .

6.5.1 Examples

The following examples are mostly in one dimension for simplicity and easy reading, but the concept applies to full three-dimensional problems. We use as a first example the generator $f = -Lp^2/2$ and, using (93), we get

$$e^{: -Lp^2/2 :} x = x - \frac{1}{2} L \underbrace{: p^2 :}_{{=-2p}} x + \frac{1}{8} L^2 \underbrace{(: p^2 :)^2}_{{=0}} x + \dots \quad (95)$$

$$= x + Lp, \quad (96)$$

$$e^{: -Lp^2/2 :} p = p - \frac{1}{2} L \underbrace{: p^2 :}_{{=0}} p + \dots \quad (97)$$

$$= p. \quad (98)$$

We have used some of the expressions derived before.

We find that Eq. (98) is the transformation of a drift space of length L .

In general, we can say that a Lie transformation acting on the phase-space coordinates x and p at a location 1 propagates them to another location 2:

$$e^{:f:} (x, p)_1 = (x, p)_2, \quad (99)$$

that is,

$$e^{:f:} x_1 = x_2, \quad (100)$$

$$e^{:f:} p_1 = p_2. \quad (101)$$

The meaning of (100) and (101) is as follows.

- i) Lie transformations describe how to go from one point to another [8, 11–13].
- ii) In our definition, it is a map from location 1 to location 2.
- iii) The motion through a machine element (drift, magnet, etc.) is fully described by f .

6.5.2 Generators of Lie transformations

We can ask the question: what is the physical meaning of f ?

- i) The generator f is related to the Hamiltonian of the element.
- ii) We use the Hamiltonian to describe the motion through an individual element.
- iii) Inside a single element the motion is ‘smooth’, while in the full machine it is not.
- iv) Given an element of length L and the Hamiltonian H , we get the generator $f = L \cdot H$.
- v) We can track ‘thick’ elements and still remain symplectic.

This forms our ‘local’ description of the objects in an accelerator, independent of other machine elements or the global behaviour.

6.5.3 Some formulas for Lie transformations

Here we provide some useful formulas using Lie transformations which are frequently used in calculations. With a being a constant, and f, g arbitrary functions of the canonical variables x and p ,

$$: a := 0, \quad e^{:a:} = 1, \quad (102)$$

$$: f : a = 0, \quad e^{:f:} a = a, \quad e^{:af:} = e^{a:f:}, \quad (103)$$

$$e^{:f:} g(x) = g(e^{:f:} x), \quad (104)$$

$$e^{:f:} G(: g :) e^{-:f:} = G(: e^{:f:} g :), \quad (105)$$

$$e^{:f:} [g, h] = [e^{:f:} g, e^{:f:} h], \quad (106)$$

$$(e^{:f:})^{-1} = e^{-:f:} \quad (107)$$

and, very importantly,

$$e^{:f:} e^{:g:} e^{-:f:} = e^{:e^{:f:} g:}. \quad (108)$$

6.5.4 Some examples for Lie transformations

For the generator

$$f = -\frac{L}{2} p^2, \quad (109)$$

using the formulas above, we obtain easily

$$e^{:f:} x = x + Lp, \quad (110)$$

$$e^{:f:} p = p. \quad (111)$$

This is the map for a drift space.

For another important generator of the form

$$f = -\frac{L}{2} (kx^2 + p^2), \quad (112)$$

we have the map

$$e^{:f:} x = e^{:-\frac{L}{2}(kx^2+p^2):} x, \quad (113)$$

$$e^{:f:} p = e^{:-\frac{L}{2}(kx^2+p^2):} p. \quad (114)$$

For its evaluation, we should remember

$$e^{:f:} g = \sum_{n=0}^{\infty} \frac{:f:^n}{n!} g. \quad (115)$$

We get (without proof)

$$e^{:-\frac{L}{2}(kx^2+p^2):} x = \sum_{n=0}^{\infty} \left(\frac{(-kL^2)^{2n}}{(2n)!} \cdot x + L \frac{(-kL^2)^{2n+1}}{(2n+1)!} \cdot p \right), \quad (116)$$

$$e^{:-\frac{L}{2}(kx^2+p^2):} p = \sum_{n=0}^{\infty} \left(\frac{(-kL^2)^{2n}}{(2n)!} \cdot p - k \frac{(-kL^2)^{2n+1}}{(2n+1)!} \cdot x \right). \quad (117)$$

The terms within the sums correspond to the trigonometric functions sine and cosine. Therefore, the generator

$$f = -\frac{L}{2} (kx^2 + p^2) \quad (118)$$

describes the map

$$e^{:f:} x = x \cos(\sqrt{k}L) + \frac{p}{\sqrt{k}} \sin(\sqrt{k}L), \quad (119)$$

$$e^{:f:} p = -\sqrt{k}x \sin(\sqrt{k}L) + p \cos(\sqrt{k}L), \quad (120)$$

which is the map for a thick, focusing quadrupole in one dimension.

6.5.5 Hamiltonians for accelerator elements

We can extend the concept to a full Hamiltonian in three dimensions and the generators for the most common magnets in three dimensions are: dipole,

$$H = -\frac{-x\delta}{\rho} + \frac{x^2}{2\rho^2} + \frac{p_x^2 + p_y^2}{2(1+\delta)}; \quad (121)$$

quadrupole,

$$H = \frac{1}{2}k_1(x^2 - y^2) + \frac{p_x^2 + p_y^2}{2(1+\delta)}; \quad (122)$$

sextupole,

$$H = \frac{1}{6}k_2(x^3 - 3xy^2) + \frac{p_x^2 + p_y^2}{2(1+\delta)}; \quad (123)$$

octupole,

$$H = \frac{1}{4}k_3(x^4 - 6x^2y^2 + y^4) + \frac{p_x^2 + p_y^2}{2(1+\delta)}. \quad (124)$$

In many cases the non-linear effects by the kinematic term are negligible ($\delta \ll 1$) and the Hamiltonian for a quadrupole

$$H = \frac{1}{2}k_1(x^2 - y^2) + \frac{p_x^2 + p_y^2}{2(1+\delta)} \quad (125)$$

can be simplified to

$$H = \frac{1}{2}k_1(x^2 - y^2) + \frac{p_x^2 + p_y^2}{2}. \quad (126)$$

In one dimensions (without y and p_y), this reduces to the previous example.

6.5.6 The use of this concept

If we know the Hamiltonian of a machine element (e.g. a magnet), then

$$e^{iLH} x_0 = x_1, \quad (127)$$

$$e^{iLH} p_0 = p_1. \quad (128)$$

This is also true for any function of x and p :

$$e^{iLH} f_0(x, p) = f_1(x, p). \quad (129)$$

Since we are now interested in the transformation of the coordinates and momenta from one location to the next, we have to multiply the Hamiltonian (invariant) with the length L of the element.

The miracles of this method are as follows.

- i) Poisson brackets always create *symplectic* maps.
- ii) The exponential form $e^{i\hbar}$ is *always* symplectic.
- iii) Better: the exponent is directly connected to the invariant of the transfer map.

6.5.7 OTM and effective Hamiltonian

We can combine many machine elements f_n by applying one transformation after the other:

$$e^{:h:} = e^{:f_1:} e^{:f_2:} \dots e^{:f_N:}. \quad (130)$$

We note the difference to the earlier treatment with linear elements:

- i) not restricted to matrices, i.e. linear elements;
- ii) we arrive at a transformation for the full ring \rightarrow a OTM;
- iii) the OTM is the exponential of the effective Hamiltonian.

It can be written as

$$\mathcal{M}_{\text{ring}} = e^{:-C\mathcal{H}_{\text{eff}}:}, \quad (131)$$

where C is the circumference of the whole ring and the effective Hamiltonian is the invariant.

6.6 Campbell–Baker–Hausdorff formula

We have to find an algorithm to combine elements in the form (130).

In the case where the two functions f and g commute (i.e. $[f, g] = [g, f]$), the concatenation is very easy:

$$e^{:h:} = e^{:f:} e^{:g:} = e^{:f+g:}. \quad (132)$$

In all other cases we can use the Campbell–Baker–Hausdorff (CBH) formula [14–17]) to obtain the effective Hamiltonian h (or OTM) and therefore

$$\begin{aligned} h = f & + g + \frac{1}{2}[f, g] + \frac{1}{12}[f, [f, g]] + \frac{1}{12}[g, [g, f]] \\ & + \frac{1}{24}[f, [g, [g, f]]] - \frac{1}{720}[g, [g, [g, [g, f]]]] \\ & - \frac{1}{720}[f, [f, [f, [f, g]]]] + \frac{1}{360}[g, [f, [f, [f, g]]]] + \dots \\ \text{or} & \\ h = f & + g + \frac{1}{2} : f : g + \frac{1}{12} : f :^2 g + \frac{1}{12} : g :^2 f \\ & + \frac{1}{24} : f : : g :^2 f - \frac{1}{720} : g :^4 f \\ & - \frac{1}{720} : f :^4 g + \frac{1}{360} : g : : f :^3 g + \dots \end{aligned} \quad (133)$$

If we proceed for all machine elements, we obtain the non-linear, effective Hamiltonian, corresponding to the OTM.

We have a special case when we combine

$$e^{:h:} = e^{:f:} e^{:g:} \quad (134)$$

if one of them (f or g) is small. Then we can truncate the series and get a very useful formula. Assume that g is small:

$$e^{:f:} e^{:g:} = e^{:h:} = \exp \left[: f + \left(\frac{: f :}{1 - e^{-:f:}} \right) g + \mathcal{O}(g^2) : \right]. \quad (135)$$

6.7 General non-linear elements

For a general thin-lens kick given by $f(x)$, we write for the Lie transformation

$$e^{:\int_0^x f(u) du:}, \quad (136)$$

which gives for the map in one dimension

$$\begin{aligned} x &= x_0, \\ p &= p_0 + f(x). \end{aligned} \quad (137)$$

For example, a thin-lens multipole of order n (i.e. $f(x) = a \cdot x^n$):

$$e^{\frac{a}{n+1} \cdot x^{n+1}} \quad (138)$$

gives for the map

$$\begin{aligned} x &= x_0, \\ p &= p_0 + ax^n. \end{aligned} \quad (139)$$

Monomials in x and p of orders n and m ($x^n p^m$)

$$e^{ax^n p^m} \quad (140)$$

give for the map (for $n \neq m$)

$$\begin{aligned} e^{ax^n p^m} x &= x \cdot [1 + a(n-m)x^{n-1}p^{m-1}]^{m/(m-n)}, \\ e^{ax^n p^m} p &= p \cdot [1 + a(n-m)x^{n-1}p^{m-1}]^{n/(n-m)}, \end{aligned} \quad (141)$$

and for the special map (for $n = m$)

$$\begin{aligned} e^{ax^n p^n} x &= x \cdot e^{-anx^{n-1}p^{n-1}}, \\ e^{ax^n p^n} p &= p \cdot e^{anx^{n-1}p^{n-1}}. \end{aligned} \quad (142)$$

6.8 From Hamiltonian to maps

We have seen that given the Hamiltonian H of a machine element of length L , the Lie operator becomes

$$f \rightarrow: f :=: LH :. \quad (143)$$

The corresponding map for an element with length L is then

$$e^{f:} = e^{:-LH:}. \quad (144)$$

Transformations derived from a Hamiltonian are always symplectic and we have (in one dimension)

$$e^{f:} x_0 = x_1, \quad (145)$$

$$e^{f:} p_0 = p_1 \quad (146)$$

or, using $Z = (x, p_x, y, p_y, \dots)$ (in two dimensions),

$$e^{f:} Z_0 = Z_1. \quad (147)$$

6.9 From maps to Hamiltonian

Assuming we do not have the Hamiltonian, but a matrix \mathcal{M} (from somewhere, including a numerical map from a tracking code), we search for the Hamiltonian. Starting from the well-known one-turn matrix

$$\mathcal{M} \equiv \begin{pmatrix} \cos(\mu) + \alpha \sin(\mu) & \beta \sin(\mu) \\ -\gamma \sin(\mu) & \cos(\mu) - \alpha \sin(\mu) \end{pmatrix} \quad (148)$$

or, rewritten,

$$\mathcal{M}Z_0 = Z_1, \quad (149)$$

we search for the corresponding form of f , such as

$$\mathcal{M} \leftrightarrow e^{:f:} \quad (= e^{: -\mu H:}). \quad (150)$$

For the linear matrix, we know that f must be a quadratic form in x and p . Any quadratic form can be written as

$$f = -\frac{1}{2}Z^* F Z \quad \left[= -\frac{1}{2}(a \cdot x^2 + b \cdot xp + c \cdot p^2) \right], \quad (151)$$

where F is a symmetric, positive-definite matrix.

Then we can write (without proof, see e.g. [8])

$$: f : Z = SFZ, \quad (152)$$

where S is the ‘symplecticity’ matrix.

Therefore, we get for the Lie transformation

$$e^{:f:} Z \leftrightarrow e^{SF} Z. \quad (153)$$

Since we have $n = 2$, we get (using the Hamilton–Cayley theorem)

$$e^{SF} = \exp \begin{pmatrix} b & c \\ -a & -b \end{pmatrix} = a_0 + a_1 \begin{pmatrix} b & c \\ -a & -b \end{pmatrix}. \quad (154)$$

We now have to find a_0 and a_1 . The eigenvalues of SF are

$$\lambda_{\pm} = \pm i \sqrt{ac - b^2}. \quad (155)$$

This tells us for the coefficients the conditions

$$e^{\lambda_+} = a_0 + a_1 \cdot \lambda_+, \quad (156)$$

$$e^{\lambda_-} = a_0 + a_1 \cdot \lambda_- \quad (157)$$

and, therefore,

$$a_0 = \cos(\sqrt{ac - b^2}), \quad (158)$$

$$a_1 = \frac{\sin(\sqrt{ac - b^2})}{\sqrt{ac - b^2}} \quad (159)$$

and

$$e^{SF} = \cos(\sqrt{ac - b^2}) + \frac{\sin(\sqrt{ac - b^2})}{\sqrt{ac - b^2}} \begin{pmatrix} b & c \\ -a & -b \end{pmatrix}. \quad (160)$$

For a general 2×2 matrix,

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}, \quad (161)$$

we get by comparison

$$\cos(\sqrt{ac - b^2}) = \frac{1}{2} \text{tr}(M) \quad (162)$$

and

$$\frac{a}{-m_{21}} = \frac{2b}{m_{11} - m_{22}} = \frac{c}{m_{12}} = \frac{\sqrt{ac - b^2}}{\sin(\sqrt{ac - b^2})} \quad (163)$$

for the quadratic form of f :

$$f = -\frac{1}{2}(a \cdot x^2 + b \cdot xp + c \cdot p^2). \quad (164)$$

6.9.1 Examples

For the example of a drift space with length L ,

$$\mathcal{M} \equiv \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix}, \quad (165)$$

we find

$$a = 0, \quad b = 0, \quad c = L \quad (166)$$

and, therefore, for the generator:

$$f = -\frac{1}{2}(Lp^2). \quad (167)$$

For the example of a thin quadrupole,

$$\mathcal{M} \equiv \begin{pmatrix} 1 & 0 \\ -\frac{1}{F} & 1 \end{pmatrix}, \quad (168)$$

we find

$$a = \frac{1}{F}, \quad b = 0, \quad c = 0 \quad (169)$$

and for the generator:

$$f = -\frac{1}{2F}(x^2). \quad (170)$$

Here I have used F for the focal length to avoid confusion with the generator f .

A very important example in one dimension is again the one-turn matrix:

$$\mathcal{M} \equiv \begin{pmatrix} \cos \mu + \alpha \sin(\mu) & \beta \sin \mu \\ -\gamma \sin \mu & \cos \mu - \alpha \sin(\mu) \end{pmatrix}. \quad (171)$$

This corresponds to

$$e^{f_2} = e^{-\mu \frac{1}{2}(\gamma x^2 + 2\alpha x p + \beta p^2)}. \quad (172)$$

In this form, f is $-\mu \cdot$ (Courant–Snyder invariant)

$$e^{f_2} = e^{-\mu \epsilon}. \quad (173)$$

We write the generator as f_2 as a special case of a transformation. We have obtained a standard Lie transformation (e^{f_2}) for the linear one-turn matrix, i.e. a rotation.

After using the normalized variables for a pure rotation,

$$\begin{pmatrix} \cos \mu + \alpha \sin(\mu) & \beta \sin \mu \\ -\gamma \sin \mu & \cos \mu - \alpha \sin(\mu) \end{pmatrix} \Rightarrow \begin{pmatrix} \cos \mu & \sin \mu \\ -\sin \mu & \cos \mu \end{pmatrix}, \quad (174)$$

$$e^{-\mu \frac{1}{2}(\gamma x^2 + 2\alpha x p + \beta p^2)} \Rightarrow e^{-\mu \frac{1}{2}(x^2 + p^2)} \quad (175)$$

and extending this to a three-dimensional linear system, we have for f_2

$$f_2 = -\frac{\mu_x}{2}(x^2 + p_x^2) - \frac{\mu_y}{2}(y^2 + p_y^2) - \frac{1}{2}\alpha_c \delta^2 \quad (176)$$

or, using the invariant action variables J ,

$$f_2 = -\mu_x J_x - \mu_y J_y - \frac{1}{2}\alpha_c \delta^2. \quad (177)$$

This is very useful in all calculations as a standard (e^{f_2}) transformation in three dimensions.

7 Non-linear normal forms

Normal-form transformations can be generalized for non-linear maps (i.e. not matrices). If \mathcal{M} is our usual OTM, we try to find a transformation:

$$\mathcal{N} = \mathcal{A} \mathcal{M} \mathcal{A}^{-1}, \quad (178)$$

where \mathcal{N} is a simple form (like the rotation we had before). Of course now we do not have matrices; we use a Lie transform with the generator F to describe the transformation \mathcal{A} :

$$\mathcal{N} = e^{-h} = \mathcal{A} \mathcal{M} \mathcal{A}^{-1} = e^{F} \mathcal{M} e^{-F}. \quad (179)$$

The canonical transformation \mathcal{A} :

$$\mathcal{N} = \mathcal{A} \mathcal{M} \mathcal{A}^{-1} \Rightarrow \mathcal{A} = e^{F} \quad (180)$$

should be the transformation to produce our simple form. The simple form should contain only the effective Hamiltonian and depend only on the action variables because it is an invariant:

- i) $h(J_x, \Psi_x, J_y, \Psi_y, z, \delta) \Rightarrow h(J_x, J_y, \delta) = h_{\text{eff}}(J_x, J_y, \delta)$;
- ii) Should work for any kind of local perturbation;

- iii) Formalism and software tools exist to find F (see e.g. [7, 22]);
- iv) Once we know $h_{\text{eff}}(J_x, J_y, \delta)$, we can derive everything.

Once we can write the map as (now in three dimensions)

$$\mathcal{N} = e^{-:h_{\text{eff}}(J_x, J_y, \delta):}, \quad (181)$$

where h_{eff} depends only on J_x, J_y and δ , then we can immediately get the tunes:

$$Q_x(J_x, J_y, \delta) = \frac{1}{2\pi} \frac{\partial h_{\text{eff}}}{\partial J_x}, \quad (182)$$

$$Q_y(J_x, J_y, \delta) = \frac{1}{2\pi} \frac{\partial h_{\text{eff}}}{\partial J_y} \quad (183)$$

and the change of path length:

$$\Delta z = -\frac{\partial h_{\text{eff}}}{\partial \delta}. \quad (184)$$

Particles with different J_x, J_y and δ have different tunes. The dependences of Q_x and Q_y on J are the amplitude detunings; the dependences on δ are the chromaticities.

7.1 Effective Hamiltonians

The effective Hamiltonian can be written in a general form (here to third order) (see e.g. [23]) as

$$\begin{aligned} h_{\text{eff}} = & \mu_x J_x + \mu_y J_y + \frac{1}{2} \alpha_c \delta^2 \\ & + c_{x1} J_x \delta + c_{y1} J_y \delta + c_3 \delta^3 \\ & + c_{xx} J_x^2 + c_{xy} J_x J_y + c_{yy} J_y^2 + c_{x2} J_x \delta^2 + c_{y2} J_y \delta^2 + c_4 \delta^4 \end{aligned} \quad (185)$$

and then (see (182) and (183))

$$Q_x(J_x, J_y, \delta) = \frac{1}{2\pi} \frac{\partial h_{\text{eff}}}{\partial J_x} = \frac{1}{2\pi} (\mu_x + 2c_{xx} J_x + c_{xy} J_y + c_{x1} \delta + c_{x2} \delta^2), \quad (186)$$

$$Q_y(J_x, J_y, \delta) = \frac{1}{2\pi} \frac{\partial h_{\text{eff}}}{\partial J_y} = \frac{1}{2\pi} (\mu_y + 2c_{yy} J_y + c_{xy} J_x + c_{y1} \delta + c_{y2} \delta^2), \quad (187)$$

$$\Delta z = -\frac{\partial h_{\text{eff}}}{\partial \delta} = \alpha_c \delta + 3c_3 \delta^2 + 4c_4 \delta^3 + c_{x1} J_x + c_{y1} J_y + 2c_{x2} J_x \delta + 2c_{y2} J_y \delta. \quad (188)$$

The physical meanings of the different coefficients we have obtained are:

- i) μ_x, μ_y : horizontal and vertical tunes;
- ii) $\frac{1}{2} \alpha_c, c_3, c_4$: linear and non-linear ‘momentum compactions’;
- iii) c_{x1}, c_{y1} : first-order chromaticities;
- iv) c_{x2}, c_{y2} : second-order chromaticities;
- v) c_{xx}, c_{xy}, c_{yy} : detuning with amplitude.

7.1.1 Example: sextupole

For the sextupole, we use the Hamiltonian (123) in one dimension. With the effective Hamiltonian we get the physical quantities of interest up to second order:

$$\mathcal{M} = e^{-\frac{\mu}{2}x^2 + p^2 + \frac{1}{2}\alpha_c\delta^2} e^{f_3} = e^{-\mu J_x + \frac{1}{2}\alpha_c\delta^2} e^{\frac{k}{6}x^3 + \frac{p^2}{2(1+\delta)}}. \quad (189)$$

We get for h_{eff} (see e.g. [6, 7])

$$h_{\text{eff}} = \mu_x J_x + \frac{1}{2}\alpha_c\delta^2 - \frac{k}{6}D^3\delta^3 - 3\frac{k}{6}\beta_x J_x D\delta \quad (190)$$

or, in three dimensions,

$$h_{\text{eff}} = \mu_x J_x + \mu_y J_y + \frac{1}{2}\alpha_c\delta^2 - kD^3\delta^3 - 3\frac{k}{6}\beta_x J_x D\delta + 3\frac{k}{6}\beta_y J_y D\delta. \quad (191)$$

When we have h_{eff} in three dimensions, we obtain (see (182) and (183))

$$Q_x(J_x, J_y, \delta) = \frac{1}{2\pi} \frac{\partial h_{\text{eff}}}{\partial J_x} = \frac{1}{2\pi} (\mu_x - 3\frac{k}{6}\beta_x D\delta), \quad (192)$$

$$Q_y(J_x, J_y, \delta) = \frac{1}{2\pi} \frac{\partial h_{\text{eff}}}{\partial J_y} = \frac{1}{2\pi} (\mu_y + 3\frac{k}{6}\beta_y D\delta) \quad (193)$$

and the change of path length

$$\Delta s = -\frac{\partial h_{\text{eff}}}{\partial \delta} = \alpha_c\delta - 3\frac{k}{6}D^3\delta^2 - 3\frac{k}{6}D(\beta_x J_x - \beta_y J_y). \quad (194)$$

With the terms $(1/2\pi)(k/2)\beta D\delta$, we have recovered the usual chromaticities from sextupoles.

7.1.2 Example: octupole

Assume a linear rotation (as always using our ‘standard map’) followed by an octupole; the Hamiltonian is 1D to keep it simple, but the formulas are valid for higher dimensions):

$$\mathcal{H} = \frac{\mu}{2}(x^2 + p^2) + k_3 \cdot \frac{x^4}{4} \quad (p = p_x), \quad (195)$$

with Hamilton’s equations leading to

$$\dot{x} = \frac{\partial \mathcal{H}}{\partial p} = \mu p, \quad (196)$$

$$\dot{p} = -\frac{\partial \mathcal{H}}{\partial x} = -\mu x - k_3 \cdot x^3. \quad (197)$$

The map, written in Lie representation, is

$$\mathcal{M} = e^{-\frac{\mu}{2}:x^2+p^2:} e^{k_3 \cdot \frac{x^4}{4}} = R e^{k_3 \cdot \frac{x^4}{4}}. \quad (198)$$

We transform it by applying

$$\mathcal{N} = A M A^{-1} = e^{F} R e^{k_3 \cdot \frac{x^4}{4}} e^{-F} = R R^{-1} e^{F} R e^{k_3 \cdot \frac{x^4}{4}} e^{-F}, \quad (199)$$

$$= R e^{R^{-1}F + k_3 \cdot \frac{x^4}{4} - F + O(\epsilon^2)} = R e^{(R^{-1}-1)F + k_3 \cdot \frac{x^4}{4} + O(\epsilon^2)}. \quad (200)$$

We have now to choose F to simplify the expression

$$(R^{-1} - 1)F + k_3 \cdot \frac{x^4}{4} \quad (201)$$

and get (e.g. [6])

$$F = -\frac{1}{64}\{-5x^4 + 3p^4 + 6x^2p^2 + x^3p \cdot (8 \cot(\mu) + 4 \cot(2\mu)) + xp^3(8 \cot(\mu) - 4 \cot(2\mu))\}. \quad (202)$$

We go back to x and p coordinates and, with

$$J = (x^2 + p^2)/2, \quad (203)$$

we can write the map to second order:

$$M = e^{-:F:} e^{:-\mu J + \frac{3}{8}k_3 \cdot J^2:} e^{:F:}. \quad (204)$$

The term $\frac{3}{8}k_3 \cdot J^2$ produces the tune shift with amplitude we know for an octupole ($< \beta^2 >$) in real space. Note that the normalized map (our most simple map)

$$R = e^{:-\mu J + \frac{3}{8}k_3 \cdot J^2:} \quad (205)$$

is again a rotation in phase space, but the rotation angle now depends on the amplitude J , as expected.

We shall look at a concrete calculation. When we have h_{eff} in one dimensions for a single octupole (see before),

$$h_{\text{eff}} = -\mu J + \frac{3}{8}k_3 \cdot J^2, \quad (206)$$

$$Q_x(J_x, J_y) = \frac{1}{2\pi} \frac{\partial h_{\text{eff}}}{\partial J_x} = -\frac{1}{2\pi} \mu_x + \frac{3}{8 \cdot 2\pi} k_3 J_x \quad (207)$$

and, with normalization in real space (i.e. $\beta \neq 1$),

$$\Delta Q_x(J_x, J_y) = \frac{3}{8 \cdot 2\pi} k_3 \langle \beta^2 \rangle J_x. \quad (208)$$

For example, for $\beta = 300$ m, $k_3 = 0.01$:

$$\Delta Q_x(J_x, J_y) = 53.7 \cdot J_x. \quad (209)$$

7.2 Example: beam-beam effects

The beam-beam interaction in a particle collider is a very localized distortion, and a very strong non-linearity in the machine. The standard perturbation theory is not sufficient to treat this problem correctly. We show here how the Lie technique is applied in this case [27]. Schematically, we show this in Fig. 12.

We are interested in the invariant h . We need the (usual) linear transform $e^{:f_2:}$ and the beam-beam interaction $e^{:F:}$, i.e.

$$e^{:f_2:} \cdot e^{:F:} = e^{:h:} \quad (210)$$

with (see before)

$$f_2 = -\frac{\mu}{2} \left(\frac{x^2}{\beta} + \beta p_x^2 \right) \quad (211)$$

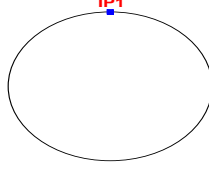


Fig. 12: Linear machine with a single interaction point (schematic)

and (see before)

$$F = \int_0^x dx' f(x'). \quad (212)$$

For a Gaussian beam, we have for $f(x)$ (see article on beam-beam effects [30])

$$f(x) = \frac{2}{x} (1 - e^{-x^2/2\sigma^2}). \quad (213)$$

As usual, we go to action-angle variables Ψ, J :

$$x = \sqrt{2J\beta} \sin \Psi, \quad p = \sqrt{\frac{2J}{\beta}} \cos \Psi \quad (214)$$

and write $F(x)$ as a Fourier series:

$$F(x) = \sum_{n=-\infty}^{\infty} c_n(J) e^{in\Psi}. \quad (215)$$

With this transformation to action-angle variables, the generator f_2 becomes very simple:

$$f_2 = -\mu J \quad (216)$$

and useful properties of Lie operators (e.g. [6, 7, 9])

$$: f_2 : g(J) = 0, \quad : f_2 : e^{in\Psi} = in\mu e^{in\Psi}, \quad g(: f_2 :) e^{in\Psi} = g(in\mu) e^{in\Psi} \quad (217)$$

and the formula (e.g. [6, 7, 9])

$$e^{:f_2:} e^{:F:} = e^{:h:} = \exp \left[: f_2 + \left(\frac{: f_2 :}{1 - e^{-:f_2:}} \right) F + \mathcal{O}(F^2) : \right] \quad (218)$$

give immediately for h

$$h = -\mu J + \sum_n c_n(J) \frac{in\mu}{1 - e^{-in\mu}} e^{in\Psi}, \quad (219)$$

$$h = -\mu J + \sum_n c_n(J) \frac{n\mu}{2 \sin(n\mu/2)} e^{(in\Psi + i\frac{n\mu}{2})}. \quad (220)$$

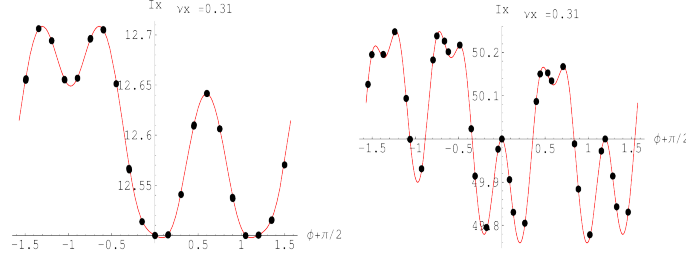


Fig. 13: Turn by turn points in phase space from simulation program, comparison with invariant (full line).

Away from resonances, the normal-form transformation gives

$$h_n = -\mu J + c_0(J) = \text{const.} \quad (221)$$

The expression

$$\frac{dc_0(J)}{dJ} \quad (222)$$

gives the detuning with amplitude due to the beam–beam interaction.

The analysis for a single interaction point gives [27]

$$h = -\mu J + \sum_n c_n(J) \frac{n\mu}{2 \sin(n\mu/2)} e^{(in\Psi + i\frac{n\mu}{2})}. \quad (223)$$

On resonance, we immediately find the conditions

$$Q = \frac{p}{n} = \frac{\mu}{2\pi} \quad (224)$$

with $c_n \neq 0$:

$$\sin\left(\frac{n\pi p}{n}\right) = \sin(p\pi) \equiv 0 \quad \text{for all integer } p \quad (225)$$

and h diverges, finding automatically all resonance conditions.

To compare with particle tracking, a simple tracking code was used [27] and the turn by turn points in phase space are shown in Fig. 13 for amplitudes of $5\sigma_x$ and $10\sigma_x$. The figure also shows the beam–beam invariant as computed from the effective Hamiltonian compared with the tracking result [27]. The agreement is remarkable.

This technique can be used to analyse and compare the strengths of non-linear effects as demonstrated in [28] and used to derive scaling laws for long range beam–beam effects in the LHC [29].

8 Differential algebra

8.1 Motivation

Analytical formulas are usually difficult to derive and often rely on simplifications and assumptions. Although they may be fast to evaluate, they can only be used as guidance to more precise calculations. It is generally accepted that elaborate tracking codes provide the most reliable description of a non-linear system. Ideally, one should like to derive global quantities such as closed orbit, tunes or detuning from the tracking codes. The tracking of a complicated system relates the output *numerically* to the input producing a numerical map and one might ask whether it is possible to relate this output to global quantities in an easy way. More precisely, can a code be used to provide input to a description or formula such that these relate more directly to global quantities? This would free us from approximate analytical calculations. They would provide ‘analytical’ formulas although they are not analytical in the classical sense, which should be considered as symbolic [6]. A typical example of such an analytical derivation would be a Taylor series derived from a tracking code.

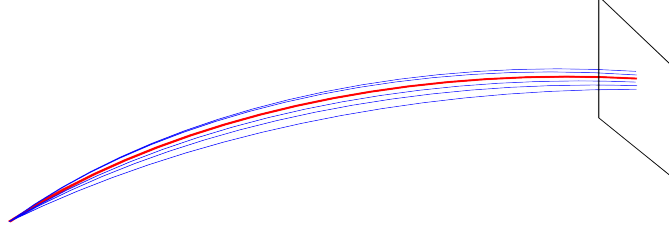


Fig. 14: Trajectories with small deviations from an ideal trajectory

8.2 Taylor series as analytical formulas

Why are Taylor series useful in the analysis of beam dynamics, in particular stability for non-linear systems? Let us study the paraxial behaviour sketched in Fig. 14. The red line indicates the ideal orbit while the blue lines are small deviations from this ideal orbit caused by distortions. If we understand how small deviations behave, we understand the system much better. Now remember the definition of the Taylor series:

$$f(a + \Delta x) = f(a) + \sum_{n=1}^{\infty} \frac{f^{(n)}(a)}{n!} \Delta x^n, \quad (226)$$

$$f(a + \Delta x) = f(a) + \frac{f'(a)}{1!} \Delta x^1 + \frac{f''(a)}{2!} \Delta x^2 + \frac{f'''(a)}{3!} \Delta x^3 + \dots \quad (227)$$

The coefficients determine the behaviour of small deviations Δx from the ideal orbit x , i.e. the Taylor expansion does a paraxial analysis of the system. If the function $f(x)$ is represented by a Taylor series:

$$f(a + \Delta x) = f(a) + \sum_{n=1}^{\infty} \frac{f^{(n)}(a)}{n!} \Delta x^n \quad (228)$$

and if it is truncated to the m th order:

$$f(a + \Delta x) = f(a) + \sum_{n=1}^m \frac{f^{(n)}(a)}{n!} \Delta x^n. \quad (229)$$

What is the relevance?

- i) There is an equivalence between the function $f(x)$ and the vector $(f(a), f'(a), f''(a), \dots, f^{(m)}(a))$.
- ii) This vector is a truncated power series algebra (TPSA) representation of $f(x)$ around a .

It can be shown that the global quantities of interest are obtained more easily and conveniently from the Taylor series map [26]. Since it is derived from a symplectic tracking code, the corresponding map is also symplectic. Our problem is how to get these coefficients without extra work.

The problem is getting the derivatives $f^{(n)}(a)$ of $f(x)$ at a :

$$f'(a) = \frac{f(a + \epsilon) - f(a)}{\epsilon}. \quad (230)$$

The problems with this approach are as follows.

- i) Need to subtract almost equal numbers and divide by small number.
- ii) For higher orders f'' , f''' , \dots , it is hopeless to obtain the required accuracy.
- iii) We can use differential algebra (DA) [24–26].

8.3 Basic concepts

The basic idea of differential algebra is the automatic differentiation of results produced by a tracking code to provide the coefficients of a Taylor series. Here we give the basic concepts of the mathematics [24–26]:

1. Define a pair (q_0, q_1) , with q_0, q_1 real numbers.

2. Define operations on a pair like

$$(q_0, q_1) + (r_0, r_1) = (q_0 + r_0, q_1 + r_1), \quad (231)$$

$$c \cdot (q_0, q_1) = (c \cdot q_0, c \cdot q_1), \quad (232)$$

$$(q_0, q_1) \cdot (r_0, r_1) = (q_0 \cdot r_0, q_0 \cdot r_1 + q_1 \cdot r_0). \quad (233)$$

3. And some ordering:

$$(q_0, q_1) < (r_0, r_1) \quad \text{if } q_0 < r_0 \text{ or } (q_0 = r_0 \text{ and } q_1 < r_1), \quad (234)$$

$$(q_0, q_1) > (r_0, r_1) \quad \text{if } q_0 > r_0 \text{ or } (q_0 = r_0 \text{ and } q_1 > r_1). \quad (235)$$

4. This implies something strange:

$$(0, 0) < (0, 1) < (r, 0) \text{ (for any positive } r), \quad (236)$$

$$(0, 1) \cdot (0, 1) = (0, 0) \rightarrow (0, 1) = \sqrt{(0, 0)}. \quad (237)$$

This means that $(0, 1)$ is between 0 and *any* real number \rightarrow infinitely small. Therefore we call this a ‘differential unit’ $d = (0, 1) = \delta$.

Of course $(q, 0)$ is just the real number q and we define the ‘real part’ and ‘differential part’ (a bit like complex numbers):

$$q_0 = \mathcal{R}(q_0, q_1) \quad \text{and} \quad q_1 = \mathcal{D}(q_0, q_1). \quad (238)$$

With our rules, we can further see that

$$(1, 0) \cdot (q_0, q_1) = (q_0, q_1), \quad (239)$$

$$(q_0, q_1)^{-1} = \left(\frac{1}{q_0}, -\frac{q_1}{q_0^2} \right). \quad (240)$$

Now we let a function f act on the pair (or vector) using our rules. For example,

$$f(x) \rightarrow f(x, 0) \quad (241)$$

acts like the function f on the real variable x :

$$f(x) = \mathcal{R}[f(x, 0)]. \quad (242)$$

What about the differential part \mathcal{D} ?

For a function $f(x)$, without proof,

$$\mathcal{D}[f(x + d)] = \mathcal{D}[f((x, 0) + (0, 1))] = \mathcal{D}[f(x, 1)] = f'(x). \quad (243)$$

The differential part of the output is the first-order derivative of the function $f(x)$. To demonstrate this, it is done most easily by looking at an example.

Assume for the function $f(x)$:

$$f(x) = x^2 + \frac{1}{x}; \quad (244)$$

then, using school calculus, we can get the derivative:

$$f'(x) = 2x - \frac{1}{x^2}. \quad (245)$$

For $x = 2$, we obtain

$$f(2) = \frac{9}{2}, f'(2) = \frac{15}{4}. \quad (246)$$

Now we proceed with the same function and use the operations defined by the differential algebra. For x in

$$f(x) = x^2 + \frac{1}{x}, \quad (247)$$

we substitute $x \rightarrow (x, 1) = (2, 1)$ and use our rules:

$$f[(2, 1)] = (2, 1)^2 + (2, 1)^{-1} \quad (248)$$

$$= (4, 4) + \left(\frac{1}{2}, -\frac{1}{4}\right) \quad (249)$$

$$= \left(\frac{9}{2}, \frac{15}{4}\right) = (f(2), f'(2)). \quad (250)$$

The computation of derivatives becomes an algebraic problem, no need for small numbers, it is exact.

8.4 Differential algebra: higher orders and more variables

8.4.1 Higher orders

The algorithm can be extended to higher order, let us say of order N :

1. the pair $(q_0, 1)$ becomes a vector of order N , $(q_0, 1) \rightarrow (q_0, 1, 0, 0, \dots, 0)$, $\delta = (0, 1, 0, 0, 0, \dots)$;
2. $(q_0, q_1, q_2, \dots, q_N) + (r_0, r_1, r_2, \dots, r_N) = (s_0, s_1, s_2, \dots, s_N)$, with $s_i = q_i + r_i$;
3. $c \cdot (q_0, q_1, q_2, \dots, q_N) = (c \cdot q_0, c \cdot q_1, c \cdot q_2, \dots, c \cdot q_N)$;
4. $(q_0, q_1, q_2, \dots, q_N) \cdot (r_0, r_1, r_2, \dots, r_N) = (s_0, s_1, s_2, \dots, s_N)$; with

$$s_i = \sum_{k=0}^i \frac{i!}{k!(i-k)!} q_k r_{i-k}. \quad (251)$$

8.4.2 Additional variables

We can go to more dimensions, i.e. more variables, with rather small modifications. If we had started with

$$x = (a, 1, 0, 0, 0, \dots), \quad (252)$$

we would get

$$f(x) = (f(a), f'(a), f''(a), f'''(a), \dots, f^{(n)}(a)). \quad (253)$$

This can be extended to more variables x, y :

$$x = (a, 1, 0, 0, 0, \dots), \quad dx = (0, 1, 0, 0, 0, \dots), \quad (254)$$

$$y = (b, 0, 1, 0, 0, \dots), \quad dy = (0, 0, 1, 0, 0, \dots) \quad (255)$$

and we get (with more complicated multiplication rules [6, 22])

$$f((x + dx), (y + dy)) = \left(f, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y}, \dots \right) (x, y). \quad (256)$$

8.5 Examples

To demonstrate the algorithm and how it can be used in practice, I show a few examples.

8.5.1 Simple example

Below is a very simple program to compute $\sin(\pi/6)$. In the first case (DATEST1) regular data types are used while in the second program (DATEST2) the usual data types are replaced by the type of a Taylor map (in the program called **my_taylor**).

```
PROGRAM DATEST1
use my_own_da
real(8) x,z,dx
my_order=3
dx=0.0
x=3.141592653/6.0 + dx
call track(x, z)
call print(z, 6)
END PROGRAM DATEST1
```

```
SUBROUTINE TRACK(a, b)
use my_own_da
real(8) a,b
b = sin(a)
END SUBROUTINE TRACK
```

```
PROGRAM DATEST2
use my_own_da
type(my_taylor) x,z,dx
my_order=3
dx=1.0.mono.1 !this is our (0,1)
x=3.141592653/6.0 + dx
call track(x, z)
call print(z, 6)
END PROGRAM DATEST2
```

```
SUBROUTINE TRACK(a, b)
use my_own_da
type(my_taylor) a,b
b = sin(a)
END SUBROUTINE TRACK
```

If we run the two programs, we obtain for the two outputs

(0,0) 0.50000000000000E+00

(0,0) 0.50000000000000E+00
(1,0) 0.8660254037844E+00
(0,1) 0.00000000000000E+00
(2,0) -0.25000000000000E+00
(0,2) 0.00000000000000E+00
(1,1) 0.00000000000000E+00
(3,0) -0.1443375672974E+00
(0,3) 0.00000000000000E+00
(2,1) 0.00000000000000E+00
(1,2) 0.00000000000000E+00

We have $\sin(\pi/6) = 0.5$ all right, but what is the rest? Let us look at

$$\sin\left(\frac{\pi}{6} + \Delta x\right) = \sin\left(\frac{\pi}{6}\right) + \cos\left(\frac{\pi}{6}\right) \Delta x^1 - \frac{1}{2} \sin\left(\frac{\pi}{6}\right) \Delta x^2 - \frac{1}{6} \cos\left(\frac{\pi}{6}\right) \Delta x^3, \quad (257)$$

i.e. we have found the coefficients of the *sine* evaluated at $\sin(\pi/6)$.

- i) We have used a simple algorithm here (sin) but it can be anything and very complex.
- ii) We can compute non-linear maps as a Taylor expansion of anything the program computes.
- iii) Simply by:
 - a) replacing regular (e.g. real) types by TPSA types (my_taylor), i.e. variables x, p are automatically replaced by $(x, 1, 0, \dots)$ and $(p, 0, 1, 0, \dots)$ etc.;
 - b) operators and functions (+, -, *, =, ..., exp, sin, ...) are automatically overloaded, i.e. behave according to the new type.

8.5.2 A FODO lattice

Another example.

- Track through a simple FODO lattice described as:
QF - DRIFT - QD
Integrate 100 steps in the quadrupoles.
Now we use three variables:
 $x, p, \Delta p = (z(1), z(2), z(3))$ (third variable needed to compute the chromaticity).

```

PROGRAM FODO1
use my_own_da
use my_analysis
type(my_taylor) z(3)
type(normalform) NORMAL
type(my_map) M,id

real(dp) L,DL,k1,k3,fix(3)
integer i,nstep

my_order=4 ! maximum order 4
fix=0.0 ! fixed point
id=1
z=fix+id

LC=62.5 ! half cell length
L=3.0 ! quadrupole length
nstep=100
DL=L/nstep
k1=0.003 ! quadrupole strength

do i=1,nstep ! track through quadrupole
z(1)=z(1)+DL/2*z(2)
z(2)=z(2)-k1*DL*z(1)/(1 + z(3))
z(1)=z(1)+DL/2*z(2)
enddo
z(1)=z(1)+LC*z(2) ! drift of half cell length

do i=1,nstep ! track through quadrupole
z(1)=z(1)+DL/2*z(2)
z(2)=z(2)-k1*DL*z(1)/(1 + z(3))
z(1)=z(1)+DL/2*z(2)
enddo
z(1)=z(1)+LC*z(2) ! drift of half cell length

call print(z(1),6)
call print(z(2),6)
M=z
NORMAL=M
write(6,*) normal%tune, normal%dtune_da
end program fodo1

```

We get the result of the program:

(0,0,0) 0.9442511679729E-01
(0,0,1) -0.9729519276183E-01

(1,0,0) 0.6972061935061E-01
(0,1,0) 0.1677727932585E+03
(1,0,1) 0.1266775134236E+01
(0,1,1)-0.3643444875882E+02
(1,0,2)-0.1603248617779E+01
(0,1,2) 0.3609522079691E+02
(1,0,3) 0.1939697138318E+01
(0,1,3)-0.3575511053483E+02

(1,0,0)-0.5300319873866E-02
(0,1,0) 0.1588490329398E+01
(1,0,1) 0.1060055415702E-01
(0,1,1)-0.5832024543075E+00
(1,0,2)-0.1590066005419E-01
(0,1,2) 0.5779004431627E+00
(1,0,3) 0.2120059477024E-01
(0,1,3)-0.5725843143370E+00

With only linear elements in the Taylor expansion,
the result for the matrix per cell is

$$\Delta x_f = 0.06972\Delta x_i + 167.77\Delta p_i, \quad (258)$$

$$\Delta p_f = -0.00530\Delta x_i + 1.5885\Delta p_i. \quad (259)$$

The outputs from the normal-form analysis are (per
cell)

Tune = 0.094425,

Chromaticity = -0.097295.

We add now to this example a thin octupole:

```

PROGRAM FODO3
use my_own_da
use my_analysis
type(my_taylor) z(3)
type(normalform) NORMAL
type(my_map) M,id

real(dp) L,DL,k1,k3,fix(3)
integer i,nstep

my_order=4 ! maximum order 4
fix=0.0 ! fixed point
id=1
z=fix+id

LC=62.5 ! half cell length
L=3.0 ! quadrupole length
nstep=100
DL=L/nstep
k1=0.003 ! quadrupole strength
k3=0.01 ! octupole strength

do i=1,nstep ! track through quadrupole
z(1)=z(1)+DL/2*z(2)
z(2)=z(2)-k1*DL*z(1)/(1 + z(3))
z(1)=z(1)+DL/2*z(2)
enddo
z(2)=z(2)-k3*z(1)**3/(1 + z(3)) ! octupole kick
z(1)=z(1)+LC*z(2) ! drift of half cell length

do i=1,nstep ! track through quadrupole
z(1)=z(1)+DL/2*z(2)
z(2)=z(2)-k1*DL*z(1)/(1 + z(3))
z(1)=z(1)+DL/2*z(2)
enddo

z(1)=z(1)+LC*z(2) ! drift of half cell length

call print(z(1),6)
call print(z(2),6)
M=z
NORMAL=M
write(6,*) normal%tune, normal%dtune_da
end program fodo3

```

The result of this program is now

```
(0,0,0) 0.9442511679729E-01
(0,0,1) -0.9729519276183E-01
```

```
(2,0,0) 0.5374370086899E+02
(0,2,0) 0.5374370086899E+02
(0,0,2) 0.1018391758451E+00
(2,0,1) 0.2035776281196E+02
.....
```

```
(1,0,0) 0.6972061935061E-01
(0,1,0) 0.1677727932585E+03
(1,0,1) 0.1266775134236E+01
(0,1,1)-0.3643444875882E+02
(3,0,0)-0.1586519461687E+01
(2,1,0)-0.1440953324752E+02
(1,2,0)-0.4362477179879E+02
```

```
.....
(1,0,0)-0.5300319873866E-02
(0,1,0) 0.1588490329398E+01
(1,0,1) 0.1060055415702E-01
(0,1,1)-0.5832024543075E+00
(3,0,0)-0.1519218878892E-01
```

Now there are non-linear elements in the Taylor expansion; the outputs from the normal-form analysis are (per cell)

Tune = 0.094425,

Chromaticity = -0.097295.

The detuning with amplitude is 53.74. This is the result we had for the analytical calculation for an octupole in the previous section.

8.5.3 Analysis using the normal form

Remember the normal-form transformation

$$\mathcal{M}\mathcal{A}\mathcal{A}^{-1} = \mathcal{R}. \quad (260)$$

The type normalform in the demonstration package also contains the maps \mathcal{A} and \mathcal{R} :

$$j2=(x^{**2}+p^{**2})*NORMAL\%A^{**}(-1)$$

Remember that $x^{**2}+p^{**2}$ is the tilted ellipse and \mathcal{A}^{-1} the transformation to the normal form, containing the optical parameters α , β , γ .

We can obtain the optical functions because:

- i) β is the coefficient of p^{**2} of invariant $j2$;
- ii) α is the coefficient of $x*p$ of invariant $j2$;
- iii) γ is the coefficient of x^{**2} of invariant $j2$.

To obtain the linear optical parameters in our code, use something like

$$\begin{aligned} \beta &= j2.sub.beta, \\ \alpha &= 0.5*j2.sub.twoalpha, \\ \gamma &= j2.sub.gamma; \end{aligned}$$

we obtain (here at the end of the cell) $\beta = 300.080714$, $\alpha = -1.358246$, $\gamma = 9.480224E-3$.

9 Summary and epilogue

In this article I have tried to show the need for a new paradigm and a new framework for the study and analysis of non-linear beam dynamics in accelerators. Emphasis was put on the use of new numerical and computational techniques which allow better (and simpler) treatment. Putting aside pedagogical issues, the dreadful use of approximations and the limitations in the standard treatments can be avoided in the era of powerful computers and computing techniques. For example, the evaluation of resonances can be easily carried beyond leading orders and brings us well ahead of standard models. Therefore, I advocate the use of contemporary tools which unify linear, non-linear, coupled, decoupled and symplectic treatments. They can be treated and handled at the same time and the theoretical framework is simplified in the humble opinion of the author. The ideas presented in this article are based on the work of numerous people and at this point I hope that the next generation of accelerator physicists will continue (or start) to follow in this direction.

Acknowledgements

I wish to thank my colleagues for their interest and many discussions on this topic. I am glad to thank the members of the SSC Central Design Group whose contributions to the topics presented here have inspired me to enter this subject. I am particularly grateful to Prof. Alex Dragt who was the first to apply the technique of Lie transformation to non-linear problems in accelerator physics. For their contributions and many discussions I appreciate the collaboration with Dr Dobrin Kaltchev and Dr Etienne Forest, the latter for the TPSA demonstration program used during the preparation of this article.

References

- [1] E. Courant and H. Snyder, *Ann. Phys.* **3** (1958) 1.
- [2] H. Wiedemann, *Particle Accelerator Physics – Basic Principles and Linear Beam Dynamics* (Springer, Berlin, 1993).
- [3] A. Chao and M. Tigner, *Handbook of Accelerator Physics and Engineering* (World Scientific, Singapore, 1998).
- [4] H. Poincaré, *Les Methods Nouvelles de la Mecanique Celeste* (Gauthier-Villars, Paris, 1892).
- [5] A.J. Lichtenberg and M.A. Lieberman, *Regular and Chaotic Dynamics* (Applied Mathematical Sciences No. 38, Springer, New York, 1983).
- [6] E. Forest, *Beam Dynamics – A New Attitude and Framework* (Harwood Academic, Amsterdam, 1998).
- [7] A. Chao, Lecture notes on topics in accelerator physics, SLAC (2001).
- [8] A. Dragt, *Lie Methods for Non-linear Dynamics with Applications to Accelerator Physics*, (University of Maryland, 2009).
- [9] W. Herr and E. Forest, Non-linear Dynamics in Accelerators, Landolt-Börnstein, New Series Vol. I/21C, *Accelerators and Colliders*, Ed. S. Myers and H. Schopper (Springer, Berlin, 2013), p.3-23].
- [10] B. Holzer, Transverse Dynamics, these proceedings (CAS 2013).
- [11] A. Dragt, *AIP Proc.* **87** (1981) 147.
- [12] A. Dragt *et al.*, *Annu. Rev. Nucl. Part. Sci.* **38** (1988) 455.
- [13] A. Dragt *et al.*, *Phys. Rev.* **A45** (1992) 2572.
- [14] F. Hausdorff, *Ber. Verh. Saechs. Akad. Wiss. (Leipz.)* **58** (1906) 19.
- [15] J. Campbell, *Proc. Lond. Math. Soc.* **28** (1897) 381; *ibid* **29** (1898) 14.
- [16] H. Poincaré, *C. R. Acad. Sci. Paris* **128** (1899) 1065; *Camb. Philos. Trans.* **18** (1899) 220.
- [17] H. Baker, *Proc. Lond. Math. Soc.* (1) **34** (1902) 347; *ibid* (1) **35** (1903) 333; *ibid* (2) **3** (1905) 24.
- [18] E. Forest and R. Ruth, *Physica* **D43** (1990) 105.

- [19] H. Yoshida, *Phys. Lett.* **A150** (1990) 262.
- [20] H. Goldstein, *Classical Mechanics* (Addison Wesley, London, 1959).
- [21] A. Wolski, Lecture notes on non-linear single particle dynamics, University of Liverpool (2006).
- [22] E. Forest, M. Berz and J. Irwin, SSC Report SSC-166 (1992).
- [23] E. Forest, D. Douglas and B. Leemann, *Nucl. Instrum. Methods* **A258** (1987) 355.
- [24] M. Berz, H.C. Hofmann and H. Wollnick, *Nucl. Instrum. Methods* **A258** (1987) 402.
- [25] M. Berz, *Part. Accel.* **24** (1989) 109.
- [26] M. Berz, *Modern Map Methods in Particle Beam Physics* (Academic Press, San Diego, 1999).
- [27] W. Herr and D. Kaltchev, Effect of phase advance between interaction points in the LHC on the beam–beam interaction, LHC Project Report 1082 (2008), unpublished.
- [28] D. Kaltchev, Analysis of long range studies in the LHC – comparison with the model, ICFA Beam–beam Workshop, CERN (2013).
- [29] R. Assmann *et al.*, Results from long range studies in the LHC – scaling with beam separation and intensity, CERN-ATS-Note-2012-070 MD (2012).
- [30] T. Pieloni, Beam–beam Effects, these proceedings (CAS 2013).